JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Chirality Measures of α-Amino Acids

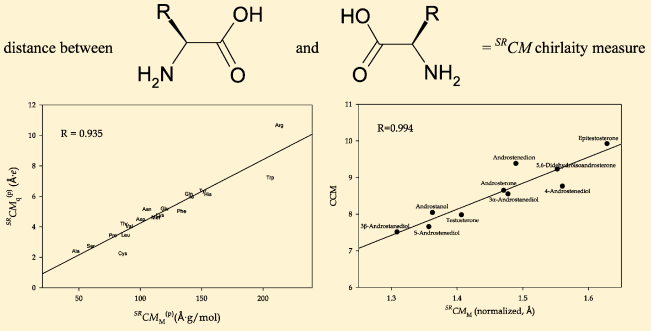
Michał H. Jamróz,[†] Joanna E. Rode,[†] Sławomir Ostrowski,[†] Piotr F. J. Lipiński,[†] and Jan Cz. Dobrowolski*[,†,‡]

[†]Industrial Chemistry Research Institute, 8 Rydygiera Street, 01-793 Warsaw, Poland

[‡]National Medicines Institute, 30/34 Chełmska Street, 00-725 Warsaw, Poland

Supporting Information

ABSTRACT: To measure molecular chirality, the molecule is treated as a finite set of points in the Euclidean R^3 space supplemented by k properties, $p_1^{(i)}$, $p_2^{(i)}$, ..., $p_k^{(i)}$ assigned to the ith atom, which constitute a point in the Property P^k space. Chirality measures are described as the distance between a molecule and its mirror image minimized over all its arbitrary orientation-preserving isometries in the $R^3 \times P^k$ Cartesian product space. Following this formalism, different chirality measures can be estimated by taking into consideration different sets of atomic properties. Here, for α-amino acid zwitterionic structures taken from the Cambridge Structural Database and for all 1684 neutral conformers of 19 biogenic α-amino acid molecules, except glycine and cystine, found at the B3LYP/6-31G** level, chirality measures have been calculated by a CHIMEA program written in this project. It is demonstrated that there is a significant correlation between the measures determined for the α-amino acid zwitterions in crystals and the neutral forms in the gas phase. Performance of the studied chirality measures with changes of the basis set and computation method was also checked. An exemplary quantitative structure−activity relationship (QSAR) application of the chirality measures was presented by an introductory model for the benchmark Cramer data set of steroidal ligands of the sex-hormone binding globulin.



INTRODUCTION

Since Louis Pasteur's discovery of "*l'hémiédrie*" (dissymmetry of the crystalline form) and his correct interpretation of the phenomenon as reflecting "*la dissymétrie des derniers éléments qui composent le crystal*" (molecular dissymmetry) in 1848,[1,2] the importance of chirality has been continuously developed. In 1874, independently, van't Hoff and Le Bell introduced the novel idea of an asymmetric carbon atom,[3,4] whereas the term "chiral" first appeared as a footnote in the Second Robert Boyle Lecture, *The Molecular Tactics of a Crystal*, delivered by Kelvin to the Oxford University Junior Scientific in 1893 and published a year later.[5,6]

The ruling IUPAC definition of chirality is as follows:[7]

> *The geometric property of a rigid object (or spatial arrangement of points or atoms) of being nonsuperposable on its mirror image; such an object has no symmetry elements of the second kind (a mirror plane, $\sigma = S_1$, a center of inversion, $i = S_2$, a rotation-reflection axis, S_{2n}). If the object is superposable on its mirror image the object is described as being achiral.*
>
> IUPAC

In other words, a figure is chiral if and only if its symmetry group does not contain any orientation-reversing isometry. However, in topology, a figure is chiral if and only if its mirror image cannot be continuously deformed into it. Thus, chirality is not only a geometric property definable by orientation-reversing isometry but also a topological property in which an isometry can be replaced by a continuous deformation.

The chirality can be measured and the measure definition depends on both space, its dimension, coordinate type, and its metric. Generally, a chirality measure should vanish for achiral structures and be greater than 0 for chiral ones. According to the review in ref 8, works on molecular chirality measures began as far back as in 1890 by Guye. In 1992, Buda, Auf der Heyde, and Mislow[9] distinguished two categories of measures (i) derived from a measure between molecule (a set) and an achiral reference and (ii) derived from a similarity measure between a molecule (a set) and its mirror image after arbitrary rotation. It is however evident that chirality measures based on chemical graph-theory form a separate class of measures since they often do not refer to geometrical properties. Several mathematical concepts have been applied for measuring chirality: symmetric difference between volumes,[10,11] distance from a symmetrical achiral set,[12−14] Hausdorff distance,[15,16] fuzzy sets,[17] graph theory,[18−20] codes of structure[21] and minimized root-mean-square distance (rms) between a set of color points and its generalization by means of probability theory.[8,21−28] Extensive reviews of chirality measures were done by Avnir, Zabrodsky, Hel-Or, and Mezey in 1998,[29] Petitjean in

Received: January 29, 2012
Published: May 15, 2012

2003,[8] and Casanova i Casas in 2006.[30] Recently, the chirality measures formulated in the frame of chemical graph theory have been reviewed by Natarajan and Basak.[31]

Even though the importance of chirality and stereochemistry is widely acknowledged in biochemistry and the stereospecificity of drug action in pharmacy, it seems that they are still not treated appropriately in quantitative structure−activity relationship (QSAR) analyses. A Pubmed literature query gives only a few articles mentioning chirality in the context of QSAR equations. On the contrary, the number of chiral drugs increases, and the pharmaceutical industry intensively introduces more and more enantiopure analogues of known racemic drugs. Chirality appears to be also underestimated in QSAR and quantitative structure−property relationship (QSPR) modeling related to physical, chemical, and technological properties of molecules. For example, in a recent comprehensive review on QSAR and QSPR by the Katrizky group, the chirality property was not mentioned at all.[32] Nevertheless, for such a fundamental property as molecular chirality, it is highly improbable that no significant QSAR representation exists.

So far, QSAR analyses of chiral substances are performed based mainly on topological indices constructed to include chirality.[33−35] However, topological indices contain information on spatial properties of molecules by virtue of an arbitrary declaration. Moreover, such a declaration is possible for an asymmetric atom easily detectable in a molecular graph, yet, hardly possible for other chirality elements or chirality defined according to the general IUPAC rule. This is why we expect that indices more closely related to 3D molecular structure may perform better. A well recognized practical method for calculation of chirality measures is the Avnir group *Continuous Chirality Measure* (CCM) accessible via the http://www.csm.huji.ac.il/ web page.[14] This method is based on distance of a chiral object from the closest abstract achiral one. So far, the physical properties of atoms and molecules were not implemented into the CCM formalism.

On the contrary, the chirality measures presented in this paper have been constructed based on molecular geometry, the IUPAC definition of chirality, and the Property Space concept. These grounds allow for introduction of different physical properties assigned to atoms and molecules to be an inherent part of a so defined chirality property. Therefore, here, the constructed measures are not only calculated for a set of α-amino acids but also tested as QSAR descriptors of biologically active molecules.

The proposed chirality measure is constructed based on the rms measure in its simple version of a minimized metric in Cartesian product of 3D Cartesian space R^3 and k-dimensional Property Space P^k, $R^3 \times P^k$. We calculated the chirality measures of the α-amino acid molecules in crystals by using the CHIMEA program written by one of the authors of this paper.[36] The study was based on the amino acid molecular structures collected in the Cambridge Structural Database (CSD)[37−41] as well as on quantum chemical calculations performed at the B3LYP/6-31G** level for 1684 individual conformers of 19 α-amino acids, except cystine and achiral glycine.

The aim of this paper is fourfold. First and foremost, we report on the construction of new chirality measures composed flexibly enough to include several physical properties of the atoms which constitute the molecule in question. Second, we present extensive calculations for chirality of α-amino acids, which are one of the most important group of chiral molecules. Third, we discuss the stability of the measures to a change of structural parameters of molecules and to a change of parameters influencing the computational accuracy. Also, the relations of the measures developed here to the chirality measures known so far for α-amino acids are shown. Fourth, the possible role of the constructed measures in QSAR modeling is presented by their inclusion into a preliminary QSAR study on sex-hormone binding globulin (SHBG) ligands.

CALCULATIONS

The conformer generation and preoptimization of the amino acid conformers at a semiempirical level were performed first by using the *Spartan 5* program,[42] and the reoptimization of all structures was performed at the B3LYP/6-31G** level by using the *Gaussian 09* suite of programs.[43] The chirality calculations were performed by using the CHIMEA program written by one of the authors of this paper,[36] assuming 50 subsequent optimization steps. For program details see below.

RESULTS AND DISCUSSION

Theory and Calculations. In everyday chemistry, chirality is mainly associated with the turning of the plane of linearly polarized light by chiral molecules and asymmetric carbon atoms. Thus, usually only a discrete chirality "measure", assigning one of the three different "values", *R*, *S*, or "none", to a chirality center is used, whereas a sequence of these symbols is assigned to the whole molecule.[44,45] Nevertheless, the chirality can be measured by more subtle methods allowing for discriminating of even slight changes in molecular chirality and asymmetry. The theory of chirality measures is very extensive.[10−31] Here, we develop measures based on similarity between enantiomers placed in Property Space, linking chirality originating from Euclidean Space with an additional physical property such as mass, charge, and label (color) of a point (atom) constituting the molecule.[8,26,29,30]

The theory applied in this study has been recently published in our paper devoted to chiral heterofullerenes.[46] Therefore, here we present only an outline of the concept without entering into the details. However here, we replaced the previously used letter χ, which might collide with the well-accepted QSAR symbol for the Kier−Hall molecular connectivity, into ${}^{SR}CM$ standing for *Sinister-Rectus (similarity) Chirality Measure*. This name is analogous to the CCM symbol of Avnir group,[14] with the SR left superscript indicating the different basis of our approach grounded in measurements of similarity between enantiomers. For a finite set $A = \{a_1, a_2, ..., a_n\}$ of n-points in R^3, the chirality measure ${}^{SR}CM(A)$ is defined as the minimum of normalized sum of all Cartesian distances between points in the set and its mirror image in reflection against a plane σ, which in an oversimplified way can be presented as

$$ {}^{SR}CM(A) = \frac{1}{a}\min\left(\sum_{i=1}^{n} w_i d_i\right) \quad (1) $$

where d_i is the distance between the ith point and a point in the mirror image translated and rotated arbitrarily, w_i is a weight of the ith distance, and a normalizes the measure. The single index for d and w labels the pairs established by a one-to-one correspondence between the points in the set and its image, whereas the minimization symbol denotes run over all permutations (of index in the mirror image) of these bijections,

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

as well arbitrary translations composed with even isometries preserving the position of a chosen point.[46]

Note that normalization guarantees the chirality measure to be independent of the w_i weight unit, thus it constitutes the *specific chirality measure*, whereas $a = 1$ makes the chirality measure proportional to the physical property defined by w_i. For a and w_i set to unit, the chirality measure reflects a sheer geometrical feature solely proportional to the number of points, whereas $a = n$ and w_i have units, the *specific geometrical chirality measure* expresses *averaged* distance between a set of points and their image. Eventually, when a is equal to the sum of r_i: the radial distances of the ith point from the geometric center, the measure is specific and unitless. The atomic mass weighted chirality measure may be normalized either by the sum of atomic masses or masses multiplied by their radial distances from mass center, while the weights themselves can be defined, for example, in analogy to the definition of reduced mass of the point and assumed mirror image.[46]

It is important that the one-to-one correspondence between the n-nondifferentiable points in the set and the mirror image can be settled in $n!$ ways; therefore, the expression (1) means that $n!$ minimized sums are considered in it. On the other hand, if the points are differentiable because of an assigned (physical) property, the number of minimizations decreases because they are running only over the points of the same property (e.g., the same mass).[46]

Here, the distance d_i used in expression (1) is generalized by assuming the properties p^1, ..., p^k to be the fourth, ..., kth coordinates and can by expressed as a Cartesian distance between two points in the $R^3 \times P^k$ space, where p_i are properties of R^3 points in a Property Space P^k (either discrete or continuous):[46]

$$d_i = \sqrt{(x_i - \overline{x}_i)^2 + (y_i - \overline{y}_i)^2 + (z_i - \overline{z}_i)^2 + (p_i^1 - \overline{p_{i_1}^1})^2 + \ldots + (p_i^k - \overline{p_{i_1}^k})^2} \quad (2)$$

Here, we consider the following properties of points (atoms): atom type (t_i), mass (m_i), and charge (q_i). The sets of atom types and masses are finite, whereas charges can change smoothly. Moreover, we assume $t_i = \overline{t}_i$, $m_i = \overline{m}_i$, and $q_i = \overline{q}_i$ the following distances ρ between the property and its image:

$$\rho(p_i - \overline{p}_i) = \begin{cases} 0 & \text{if } p_i = \overline{p}_i \\ \infty & \text{otherwise} \end{cases}, \text{ where } p_i = t_i, m_i, q_i \quad (3)$$

for which an additional unit converting coefficients has not been defined because for both values they are meaningless. Definition (3) guarantees that the minimization condition can be satisfied if, and only if, the proper atoms are paired by mirroring.

Another important issue comes from a remark that in a crystal a molecule may occupy two or more nonequivalent positions and may occur in different forms. In such a case, the chirality of a molecule A in a crystal phase K, ${}^{SR}CM(A(K))$ can be understood as the value averaged over all k-nonequivalent structures of A in the elementary cell:

$${}^{SR}CM(A(\mathrm{K})) = \frac{1}{k}\sum_{i=1}^{k} l_i {}^{SR}CM(A_i) \quad (4)$$

where l_i is number of molecule A_i in the elementary cell.

Moreover, a flexible molecule A in Γ—the gas phase; rare-gas-low-temperature matrices; polymer matrices; or in dissolved state—exists in several conformations of substantially different chirality measures. In such situations, the chirality of a molecule, ${}^{SR}CM(A(\Gamma))$, can be determined as the value averaged over all conformers and weighted by Boltzmann populations of the conformers b_i:

$${}^{SR}CM(A(\Gamma)) = \sum_{i=1}^{k} l_i {}^{SR}CM(A_i) \quad (5)$$

where the Boltzmann factor for the conformer A_i is $(1/Z)\exp((E_i - E_{\min})/RT)$ and E_i is energy (Gibbs free energy) of the ith conformer, $E_{\min}$ is energy of minimum energy conformer, R is the gas constant, T is the temperature, and $Z = \sum_{i=1}^{k} \exp((E_i - E_{\min})/RT)$.

Finally, let us stress that for non rigid molecules adopting many different conformation a serious problem may arise. Indeed, most of the conformers of an achiral molecule adopts conformations that are chiral. In agreement with the mathematical definition, a measure is a positively defined functional which to both enantiomers should assign the same positive value. Thus, chirality measures of the two conformational enantiomers of an achiral molecule contribute to the molecular chirality with the same positive value weighted by a population factor. However, for an achiral nonrigid molecule it will always be a conformation exhibiting an orientation reversing symmetry element and for it the measure will be zero. No such conformation can exist for a chiral molecule regardless of its conformation. Thus, inspection of the chirality values to check whether some values are vanishing or not is a necessary step in an automatic chirality analysis, unless sheer conformational behavior of an achiral molecule is not a primary aim of a study.

The above concept is implemented into the CHIMEA program, calculating the geometric, labeled, mass, and charge chirality measures by using a modification of the distance (2) in the following form:

$$d_i^w = \sqrt{(w_i x_i - \overline{w_i x_i})^2 + (w_i y_i - \overline{w_i y_i})^2 + (w_i z_i - \overline{w_i z_i})^2 + (p_i - \overline{p}_i)^2} \quad (2a)$$

which reduces to $w_i d_i$ when $w_i = \overline{w_i}$ but is slightly different if we allow comparison of different properties. In expression (2a), w denotes mass or charge of the ith atom and the ith atom in the image, whereas p_i may also point connectivity up to 2nd and 4th order.

To calculate the measures, the CHIMEA program places specifically the molecule and its mirror image, measures an introductory (geometric (pure), labeled, mass, or charge) measure value which next is iteratively minimized in repetitive alternation of translations and rotations of the molecules. In search of the measure, several local and global minima can be found. To increase the possibility for finding the global minimum, several starting arrangements of the image against the molecule are considered. This is especially important for "ball-like" structures such as heterofullerenes,[46] for which the Monte Carlo routine additionally generates several hundred hazardous rotations. It is very important that the different properties and weights lead to different arrangements of enantiomers in the minimized chirality distance (Figure 1).

***Sinister-Rectus* Chirality Measures of α-Amino Acids.** *Chirality of α-Amino Acid Zwitterions in the Crystal Phase.* The α-amino acids and carbohydrates are probably the most important chiral molecules. However, from the point of view of determination of chirality measures, these fundamental biomolecules are quite difficult and challenging. Indeed, each of them can adopt several conformations which are not easily

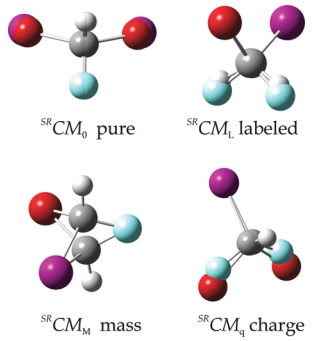


Figure 1. Minimized arrangement of enantiomers of the CFClBr molecule obtained by using four different weights in minimization leading to determine the ^{SR}CM chirality measure.

determined, neither experimentally nor computationally. Therefore, the chirality measure of amino acid molecules in all phases, except the solid phase, is a result of chirality measures of all conformers and the function of their population in given conditions (eq 5). As a consequence of the computational complexity of the chirality measure determination in gas and liquid phases, we first determine the measures for the most univocal conditions offered by the structures of amino acid molecules frozen in crystals. The appropriate data is accessible in the Cambridge Structural Database (CSD).[47−116] It is important that in crystals, amino acids exist exclusively in zwitterionic forms and occupy not more than two inequivalent positions. The chirality measures were calculated only for those amino acid molecules for which all positions of H atoms were (directly or indirectly) determined (Table 1). Each of the measures was taken usually after a minimum of 50 optimization cycles, in which each cycle tested several thousands of comparisons of distance between molecule and the image.

To properly understand the values of chirality measures listed in Table 1, it is important to remember that they are the upper estimates of the true measures. The calculations always yield the lowest value found after a finite number n of optimization cycles. In consequence, the measure found in the $(n + 1)$ th optimization cycle cannot be greater that found in the nth one.

$$^{SR}CM^{i}_{calc}(n) \geq {}^{SR}CM^{i}_{calc}(n+1) \geq {}^{SR}CM^{i}(\infty) = {}^{SR}CM^{i}_{true}$$

where i runs over the chirality measure type considered in this paper.

However, it is not impossible that for a certain entry the value will be somewhat greater than the true value, yet, this is an unavoidable property of the type of optimization algorithm used. A deeper insight into the values found (Table 1) shows that for a given α-amino acid

$$^{SR}CM_0 < {}^{SR}CM_L < {}^{SR}CM_M$$

Also, for two amino acids, A1 and A2, the chirality measures are different:

$$^{SR}CM_i^{A1} \neq {}^{SR}CM_i^{A2}$$

In most cases, the chirality measures of the same amino acid crystallizing in the same space group, but measured in different laboratories with different accuracy and conditions, differs but only slightly. For example, for alanine structures determined between 1966 and 2005, the $^{SR}CM_0$, $^{SR}CM_L$, and $^{SR}CM_M$ values vary from 3.8 to 4.1 Å, from 10.6 to 11.7 Å, and from 58.5 to 63.1 Å·g/mol, respectively. Note, that for the analyzed structures, the refinement factor (R-factor) varied from 1.59 to 8.44 and the determined molar volume differed substantially as well (Table 1). As expected, the measures for D-alanine are practically the same as for L-alanine (Table 1). Indeed, the distance between each of them and a mirror image should always be the same, if they were measured for the same space group and conditions. However, if needed, the assignment of the sign connected with R/S (D/L or $\pm$) isomerism can be introduced at the ground of an additional convention.

In the case of serine, we can compare values for as much as 26 structures. The $^{SR}CM_0$, $^{SR}CM_L$, and $^{SR}CM_M$ values vary from 5.3 to 6.8 Å, from 13.6 to 16.3 Å, and from 74.8 to 96.7 Å·g/mol, respectively. However, the lowest $^{SR}CM_0$ values were determined only for three structures, the highest values only for two of them, and for the majority, the values were focused around 6 Å. It is difficult to judge the cause of the observed spread of chirality measure values; however, it should be remembered that the position of the H-atoms is usually determined by additional calculation followed the structure determination. The routines leading to assignment of H-atom positions depend on the quality of the determined structure and several additional assumptions. Additionally, for the set of analyzed structures, the quality of crystals differed and the structures were determined under different conditions.

D-Alloleucine is an interesting case: it crystallizes in the $P2_1$ space group with two inequivalent molecules in the elementary cell.[59,60] However, the $^{SR}CM_0$ measures of those two molecules do not differ, while the $^{SR}CM_L$ and $^{SR}CM_M$ measures of the two molecules are equal to ca. 20 and 30 Å and ca. 75 and 130 Å·g/mol, respectively. This suggests a great sensitivity of the $^{SR}CM_L$ and $^{SR}CM_M$ measures to conformational changes. A similar behavior was found for L-izoleucine.[57] On the other hand, for DL-histidine crystallizing in the $P2_1/c$ space group, $^{SR}CM_0$ is equal to 16.6 Å, while for L-histidine crystallizing in either the $P2_1$ or $P2_12_12_1$ group, $^{SR}CM_0$ is equal to 10.5 Å.[110−115] Also the $^{SR}CM_L$ and $^{SR}CM_M$ measures differ: they are equal to 24 and 15 Å and 115 and 100 Å·g/mol, respectively.

Note, that although different weights lead to different arrangements of enantiomers in minimal geometries (Figure 1), and this can occur for any of the conformers independently, the measures for α-amino acid zwitterions in the crystal phases intercorrelate (Figure 2). The correlations are not strong: R = 0.795, 0.775, and 0.857, for the plot of $^{SR}CM_0^{(c)}$ vs $^{SR}CM_L^{(c)}$, $^{SR}CM_0^{(c)}$ vs $^{SR}CM_M^{(c)}$, and $^{SR}CM_L^{(c)}$ vs $^{SR}CM_M^{(c)}$, respectively, where (c) denotes averaging over all crystal data available for the given amino acid (Table 1).

Several factors influence the quality of such correlations. First, each assumed weight and precise definition of a given chirality measure, leads to different minimized mutual arrangements of the two enantiomers. Second, the quality of the crystal structure obtained from different measurements varied (Table 1), and the value taken for correlation was an unweighted average over all of these structures. Third, for some amino acids over 20 different crystals were measured with acceptable quality (serine), whereas for others only one such structure (asparagine) was found in CSD (Table 1). Fourthly, the number of local minima on a hypersurface, which ought to be scanned in search of the global chirality measure minimum, differ with the measure definition.

Indeed, the dependence of number of local minima on the chirality measure hypersurface (CMH) may be crucial for quality of correlations. In the case of pure chirality measure any

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

Table 1. Chirality Measures, $^{SR}CM_0$ (Å), $^{SR}CM_L$ (Å), and $^{SR}CM_M$ (Å·g/mol) for Zwitterionic Amino Acid Molecular Structures Collected in the Cambridge Structural Database (CSD)[a]

molecule	$^{SR}CM_0$	$^{SR}CM_L$	$^{SR}CM_M$		R-factor	mol vol	space group	ref year
alanine				C3 H7 N O2				
ALUCAL04	4.0	11.6	60.0	D-alanine	8.44	431.212	$P2_12_12_1$	42 2005
ALUCAL05	4.0	11.6	60.0	D-alanine	7.99	421.465	$P2_12_12_1$	42 2005
DLANIN01	3.8	10.7	58.0	DL-alanine	2.52	422.858	$Pna2_1$	43 2001
LANIN	3.8	10.6	58.9	L-alanine	4.9	430.636	$P2_12_12_1$	44 1966
LANIN03	4.0	11.4	60.2	L-alanine	6.1	430.636	$P2_12_12_1$	45 1966
LANIN12	4.0	11.5	59.3	L-alanine	1.59	421.067	$P2_12_12_1$	46 1988
LANIN22	4.0	11.2	59.3	L-alanine	2.2	429.4	$P2_12_12_1$	47 1970
LANIN23	4.1	11.6	60.2	L-alanine	5.88	423.301	$P2_12_12_1$	42 2005
alloisoleucinee				C6 H13 N O2				
DAILEU01-1	13.9	29.9	129.4	D-alloisoleucine	4.83	710.053	$P2_1$	54 2000
DAILEU01-2	14.4	20.5	74.7	D-alloisoleucine	4.83	710.053	$P2_1$	54 2000
DAILEU-1	14.7	29.9	130.6	D-alloisoleucine	11.8	733.476	$P2_1$	55 1975
DAILEU-2	11.6	17.5	70.0	D-alloisoleucine	11.8	733.476	$P2_1$	55 1975
aspargine				C4 H8 N2 O3				
VIKKEG	13.9	20.9	108.0	L-asparagine	3.24	273.059	$P2_1$	103 2007
aspartic acid				C4 H7 N O4				
DLASPA02	10.9	14.2	83.5	DL-aspartic acid	4.2	1073.19	$C2/c$	89 1989
DLASPA03	10.5	13.3	82.9	DL-aspartic acid	3.13	1053.45	$C2/c$	90 1998
DLASPA10	10.8	13.6	83.1	DL-aspartic acid	3.1	1075.43	$C2/c$	91 1973
DLASPA11	10.9	13.6	83.0	DL-aspartic acid	3.85	1074.86	$C2/c$	92 2007
LASPRT	11.2	17.7	105.7	L-aspartic acid	2	269.438	$P2_1$	93 1968
LASPRT02	10.7	18.2	106.1	L-aspartic acid	2.73	269.09	$P2_1$	94 2006
LASPRT03	11.0	18.3	109.1	L-Aspartic acid	1.9	263.488	$P2_1$	95 2007
LASPRT04	11.0	18.0	109.0	L-aspartic acid	2.55	263.488	$P2_1$	95 2007
cysteine				C3 H7 N1 O2 S1				
BOQCUF	8.2	11.6	50.3	DL-cysteine	6	558.437	$P2_1/a$	69 1999
LCYSTN04-1	7.5	14.9	76.4	L-cysteine	3.11	528.473	$P2_1$	70 1996
LCYSTN04-2	6.5	17.9	76.8	L-cysteine	3.11	528.473	$P2_1$	70 1996
LCYSTN22	6.2	17.2	61.2	L-cysteine	1.7	526.442	$P2_12_12_1$	71 2005
LCYSTN24	9.2	14.4	56.6	L-cysteine	5.7	451.785	$P2_12_12_1$	72 2006
LCYSTN25	8.8	14.3	57.7	L-cysteine	4.02	435.197	$P2_12_12_1$	72 2006
cystine				C6 H12 N2 O4 S2				
LCYSTI10	23.0	38.4	267.5	L-cystine	12.3	1432.73	$P6_122$	73 1959
LCYSTI11	20.9	42.0	309.9	L-cystine	9.7	978.374	$P4_1$	74 1974
LCYSTI14	20.0	39.7	267.4	L-cystine	1.4	1419.36	$P6_122$	75 1999
LCYSTI15	20.8	38.6	264.1	L-cystine	4.76	1424.33	$P6_122$	76 2005
LCYSTI16	20.9	38.4	263.6	L-cystine	3.51	1415.35	$P6_122$	76 2005
LCYSTI17	20.9	38.4	261.3	L-cystine	3.56	1366.71	$P6_122$	76 2005
LCYSTI18	21.8	38.5	259.5	L-cystine	4.83	1327.75	$P6_122$	76 2005
LCYSTI19	21.8	38.4	255.0	L-cystine	9.34	1286.15	$P6_122$	76 2005
glutamic acid				C5 H9 N O4				
LGLUAC01	13.4	23.6	123.4	L-glutamic acid	6	618.048	$P2_12_12_1$	96 1998
LGLUAC02	15.7	24.0	148.1	L-glutamic acid	3.4	635.944	$P2_12_12_1$	97 1980
LGLUAC03	15.4	23.9	148.8	L-glutamic acid	2.1	637.998	$P2_12_12_1$	98 1980
LGLUAC11	13.7	25.4	124.2	L-glutamic acid	2.6	620.114	$P2_12_12_1$	99 1972
YUYMOU	10.6	17.4	141.9	DL-glutamic acid	3.8	613.339	$P2_1/n$	111 1995
glutamine				C5 H10 N2 O3				
GLUTAM01	7.2	23.3	142.3	L-glutamine	3.2	636.534	$P2_12_12_1$	100 1973
GLUTAM02	7.1	23.3	143.7	L-glutamine	1.4	633.415	$P2_12_12_1$	101 2001
TACQUJ	11.7	15.8	106.6	DL-glutamine	4.06	711.789	$P2_1/c$	102 1996
histidine				C6 H9 N3 O2				
DLHIST	16.5	25.2	118.3	DL-histidine	13	677.87	$P2_1/c$	105 1974
DLHIST01	16.7	23.6	115.6	DL-histidine	2.96	664.77	$P2_1/c$	106 1999
LHISTD02	10.4	14.7	100.0	L-(+)-histidine	10	358.99	$P2_1$	107 1972
LHISTD04	10.8	14.9	97.9	L-histidine	3.9	357.442	$P2_1$	108 1993
LHISTD10	11.1	15.1	99.6	L-(+)-histidine	3.4	715.286	$P2_12_12_1$	109 1972
LHISTD13	10.6	15.3	99.7	L-histidine	5.8	709.784	$P2_12_12_1$	110 1972

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

Table 1. continued

molecule	$^{SR}CM_0$	$^{SR}CM_L$	$^{SR}CM_M$		R-factor	mol vol	space group	ref year
isoleucine				C6 H13 N O2				
LISLEU02-1	14.0	18.6	82.8	L-isoleucine	5.24	712.072	$P2_1$	56 1996
LISLEU02-2	13.4	30.3	137.8	L-isoleucine	5.24	712.072	$P2_1$	56 1996
XADVED-1	14.0	30.2	133.4	L-isoleucine	3.26	351.42	$P1$	54 2000
XADVED-2	12.7	31.5	142.6	L-isoleucine	3.26	351.42	$P1$	54 2000
leucine				C6 H13 N O2				
DLLEUC	13.0	32.4	140.7	DL-leucine	5.8	365.688	$\bar{P}1$	57 1975
LEUCIN01-1	12.4	19.1	103.8	L-leucine	5.8	748.171	$P2_1$	58 1986
LEUCIN01-2	11.5	19.3	97.1	L-leucine	5.8	748.171	$P2_1$	58 1986
LEUCIN02-1	13.6	18.0	100.0	L-leucine	4.35	733.965	$P2_1$	59 1996
LEUCIN02-2	13.9	18.8	104.1	L-leucine	4.35	733.965	$P2_1$	59 1996
LEUCIN03	12.8	16.5	81.2	L-*a*-leucine	3.31	742.535	C_2	60 2007
phenylalanine				C9 H11 N O2				
SIMPEJ-1	8.2	12.3	97.7	(*R*)-phenylalanine	14.72	1667.6	C_2	61 1990
SIMPEJ-2	8.2	12.6	106.4	(*R*)-phenylalanine	14.72	1667.6	C_2	61 1990
proline				C5 H9 N O2				
PROLIN	8.1	16.8	102.3	L-proline	16.9	541.741	$P2_12_12_1$	62 1965
QANRUT	5.9	20.1	104.5	DL-proline	3.95	542.699	$P2_1/c$	63 2005
QANRUT01	6.4	19.6	103.1	DL-proline	8.4	524.397	$P2_1/c$	64 2006
serine				C3 H7 N O3				
DLSERN	5.7	15.9	83.2	DL-serine	14.5	453.406	$P2_1/a$	77 1953
DLSERN02	5.8	15.7	82.5	DL-serine	3.2	455.198	$P2_1/a$	78 1974
DLSERN11	5.6	16.3	82.8	DL-serine	2	453.964	$P2_1/a$	79 1973
DLSERN12	5.7	16.2	83.4	DL-serine	3.09	452.362	$P2_1/a$	53 2002
DLSERN13	5.6	16.2	83.1	DL-serine	2.23	454.825	$P2_1/a$	80 2005
DLSERN14	5.6	16.1	83.1	DL-serine	2.28	452.827	$P2_1/a$	80 2005
DLSERN15	5.8	16.2	83.4	DL-serine	1.96	453.134	$P2_1/a$	80 2005
DLSERN16	5.8	15.5	82.2	DL-serine	5.56	455.848	$P2_1/n$	81 2006
DLSERN17	5.9	15.5	92.5	DL-serine	4.49	454.107	$P2_1/n$	81 2006
DLSERN18	5.9	15.4	82.5	DL-serine	4.23	453.613	$P2_1/n$	81 2006
DLSERN19	5.3	15.3	86.1	DL-serine	3.61	365.102	$P2_1/n$	81 2006
LSERIN01	5.9	14.8	78.2	L-(−)-serine	5.9	451.594	$P2_12_12_1$	78 1974
LSERIN10	5.9	14.8	78.7	L-(−)-serine	4.4	448.776	$P2_12_12_1$	82 1973
LSERIN11	5.9	15.2	78.6	L-serine	8.27	436.485	$P2_12_12_1$	83 2005
LSERIN12	5.3	15.2	77.5	L-serine	7.33	419.181	$P2_12_12_1$	83 2005
LSERIN13	5.0	15.2	77.6	L-serine	6.49	404.488	$P2_12_12_1$	83 2005
LSERIN14	5.2	15.3	77.0	L-serine	6.64	395.75	$P2_12_12_1$	83 2005
LSERIN15	4.7	15.1	78.1	L-serine	6.03	390.09	$P2_12_12_1$	83 2005
LSERIN16	5.9	14.2	75.6	L-serine	4.82	374.198	$P2_12_12_1$	83 2005
LSERIN18	6.0	14.9	78.1	L-serine	5.03	450.582	$P2_12_12_1$	84 2005
LSERIN19	5.3	14.8	77.5	L-serine	4.96	446.85	$P2_12_12_1$	84 2005
LSERIN20	6.1	14.7	77.1	L-serine	3.23	443.475	$P2_12_12_1$	84 2005
LSERIN21	5.3	14.0	76.1	L-serine	4.55	350.689	$P2_12_12_1$	85 2006
LSERIN26	5.7	15.4	75.7	L-serine	5.54	392.807	$P2_12_12_1$	86 2006
LSERIN27	5.6	14.2	74.6	L-serine	5.26	376.661	$P2_12_12_1$	86 2006
LSERIN28	5.7	13.5	74.5	L-serine	3.72	355.046	$P2_12_12_1$	86 2006
threonine				C4 H9 N O3				
LTHREO01	10.3	14.0	77.2	L-threonine	6.8	545.486	$P2_12_12_1$	87 1973
LTHREO02	10.4	14.7	77.9	L-threonine	6.8	545.486	$P2_12_12_1$	87 1973
LTHREO03	10.3	13.7	76.7	L-threonine	6.2	530.511	$P2_12_12_1$	88 1997
thyrosine				C9 H11 N O3				
DLTYRS	11.8	25.6	175.6	DL-tyrosine	3.7	837.879	$Pna2_1$	65 1973
FAZHET01	11.6	26.5	167.8	D-tyrosine	3.09	839.685	$P2_12_12_1$	66 2001
LTYROS10	11.4	23.8	171.6	L-tyrosine	4.9	850.888	$P2_12_12_1$	67 1972
LTYROS11	11.9	26.5	198.6	L-tyrosine	4	851.406	$P2_12_12_1$	68 1973
tryptophan				C11 H12 N2 O2				
QQQBTP02	20.1	29.5	239.4	DL-tryptophan	4.03	989.385	$P2_1/c$	104 2004
valine				C5 H11 N O2				
AHEJEC01-1	10.6	13.7	82.5	D-valine	3.23	605.113	$P2_1$	48 2002
AHEJEC01-2	5.2	15.3	66.7	D-valine	3.23	605.113	$P2_1$	48 2002
AHEJEC02-1	5.1	15.3	66.2	D-valine	3.79	612.592	$P2_1$	48 2002

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

Table 1. continued

molecule	$^{SR}CM_0$	$^{SR}CM_L$	$^{SR}CM_M$		R-factor	mol vol	space group	ref year
valine				C5 H11 N O2				
AHEJEC02-2	10.5	13.6	81.6	D-valine	3.79	612.592	$P2_1$	48 2002
AHEJEC03-1	10.4	13.6	81.6	D-valine	3.25	614.078	$P2_1$	48 2002
AHEJEC03-2	5.2	15.2	65.7	D-valine	3.25	614.078	$P2_1$	48 2002
AHEJEC-1	10.5	13.8	82.0	D-valine	3.61	608.886	$P2_1$	48 2002
AHEJEC-2	5.7	15.4	66.7	D-valine	3.61	608.886	$P2_1$	48 2002
BERQAQ-1	6.6	16.6	76.8	L-valine	4	577.602	$P2_1$	49 1999
BERQAQ-2	10.2	13.8	48.0	D-2-aminobutanoic acid C4 H9 N O2	4	577.602	$P2_1$	49 1999
LVALIN01-1	12.2	13.6	82.8	L-valine	3.4	606.031	$P2_1$	50 1996
LVALIN01-2	9.6	14.9	66.8	L-valine	3.4	606.031	$P2_1$	50 1996
VALIDL	10.3	13.6	76.9	DL-valine	10.1	588.264	$P2_1/c$	51 1969
VALIDL02	10.5	13.6	77.5	DL-valine	4.52	287.088	$\bar{P}1$	52 1996
VALIDL03	10.2	14.0	78.1	DL-valine	2.95	287.816	$\bar{P}1$	53 2002

[a]Only amino acid crystals with determined positions of all H-atom were analyzed. The first column lists the CSD crystal code, and the last four columns give information on structure solvation reliability factor, molecular volume (Å^3), space group, and publication date which differentiates the structures. If two inequivalent molecular structures were present in the cell unit, the single molecules are signed by numbers 1 and 2 added to the crystal code.

pair of atoms in a molecule and its mirror image can be compared, yielding number of $n!$ permutations, whereas for the labeled measure this number decreases to $i!{\cdot}j!{\cdot}...{\cdot}k!$, where $(i + j + ... + l = n)$ and $i, j, ..., k$ are the numbers of atoms in the molecule labeled by the same symbol. The same can be said concerning the number of searches of mass global minimum. However, an additional condition is connected to the mass weights which "attract" the search to the minima related to large mass more than the other. Thus the spread of points of pure measure is much greater for the combinatorial reasons, while the spread of the mass measures is much lower. This was the main reason why in our study on chirality measures of chiral diheterofullerenes we concluded that "pure geometrical chirality measure should not be used in QSAR predictions, while the labeled and mass weighted measures are promising QSAR descriptors of the molecular chirality."[46] We think that the lower dimension of the CMH in the case of $^{SR}CM_L^{(c)}$ and $^{SR}CM_M^{(c)}$ measures is the direct cause of the correlation between these measures being the strongest (Figure 2c).

Chirality of α-Amino Acids in the Gas Phase. The chirality measures of 19 α-amino acids in the gas phase (298 K, 1 atm) were estimated based on structures calculated at the B3LYP/6−21G** level. Glycine and cystine were omitted, glycine because of its achirality and cystine because of a very large number of cystine conformers ($2^4{\cdot}3^9 = 78\,732$) necessary to be considered in the systematic search of the conformers. The choice of the computational method and the basis set was determined by the great number of conformations adopted by amino acid molecules (Table 2). The most rigid amino acid, proline, has only three conformers stable at the chosen level of theory, whereas for the most flexible one, arginine, we found as many as 520 stable conformers. Meticulous elimination of the repetition of the calculated structures reduced the overall number of structures from over 2200 to 1684.

The chirality measures, $^{SR}CM_0^{(p)}$, $^{SR}CM_L^{(p)}$, $^{SR}CM_M^{(p)}$, and $^{SR}CM_q^{(p)}$, where (p) denotes population averaging, for all conformers of the calculated α-amino acid structures are gathered in Tables 1SI−19SI of the Supporting Information. All four measures vary substantially with the conformation. For example for the most populated alanine conformer, the values $^{SR}CM_0$, $^{SR}CM_L$, $^{SR}CM_M$, and $^{SR}CM_q$ are equal to ca. 7.7 Å, 11.5 Å, 54.8 Å·g/mol, and 2.6 Å·e, respectively (Table 1SI), while the range of changeability for these values is 5.5−8.7 Å, 8.1−15.0 Å, 26.7−77.6 Å·g/mol, and 1.5−4.1 Å·e, respectively. Similar can be noted for all other amino acid chirality measures: the value averaged over the whole population is very different from the values for the particular conformers (Tables 1SI−19SI).

Again, despite the fact that different weights lead to different arrangements of enantiomers in minimal geometries (Figure 1), and this occurs for each conformer independently, the measures of α-amino acids in the gas phase intercorrelate (Figure 3). Unexpectedly, the correlation is the strongest ($R = 0.951$) for plot of $^{SR}CM_0^{(p)}$ vs $^{SR}CM_L^{(p)}$ (Figure 3A), yet the other correlations between pure chirality measure and mass or charge are the weakest. Correlations between the labeled measure and mass and charge as well as between mass and charge exhibit a correlation coefficient exceeding 0.920 (Figure 3D−F). The last conclusion is in agreement with the discussion noted above for the correlations between the measures for the amino acid zwitterions. However, at the moment, we do not understand the high quality of the correlation between pure and labeled measures (Figure 3A).

Interestingly, there are weak, yet significant, correlations between chirality measures of amino acids in the gas phase and analogous measures for amino acid zwitterions in the crystal phase (Figure 4). It is not obvious why such correlations should be statistically significant. First, amino acids in crystals occur as zwitterions, whereas in the gas phase are neutral. Second, and more importantly, in crystals they are frozen in one or two nonequivalent conformations constrained by lattice forces (Table 1), whereas in gas they freely adopt multitudes of conformations (520 for arginine), which in our approach are averaged over entire populations. Thus, one might expect that for a number of amino acids the conformation of zwitterions in crystal can be quite different from the population averaged conformation of the gas phase neutral form. Despite the above concerns, the quality of the obtained correlations (Figure 4) may be a consequence of all the factors discussed in previous paragraphs for the correlations between the measures of zwitterions themselves. Taking into account these arguments, we think that the quality of correlations is quite satisfying. This strongly suggests that the chirality measures obtained, based on calculations, can be valuable parameters in QSAR analyses for

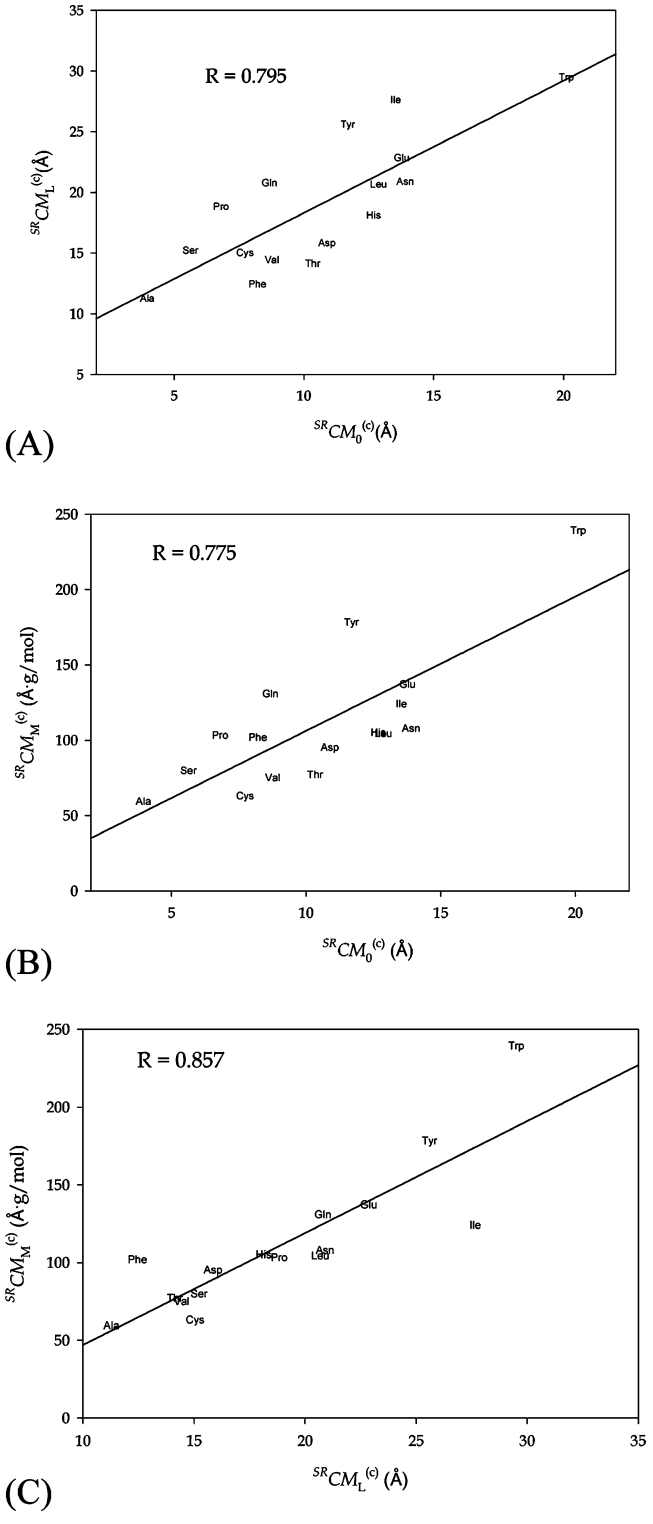


Figure 2. Mutual linear correlations between chirality measures calculated for zwitterionic forms of amino acids in crystals (averaged over all available data for the given crystal). $^{SR}CM_0$, $^{SR}CM_L$, and $^{SR}CM_M$ denote the pure, labeled, and mass chirality measure, and (c) denotes that the values were averaged over all data available for the crystals of a given amino acid.

molecules operating in different phases of condensation and ordering.

At this moment, a methodological question arises: to what extent are the measures (i) reproducible, (ii) sensitive to basis set changes, and (iii) sensitive to computational method change? This is an important question because it is known that many experimental parameters require a high computational level to fairly reproduce the experimental data for amino acids.[117−123] To answer this question we recalculated the four measures for 50 cysteine conformers using: (i) the same B3LYP/6-31G** method, (ii) the B3LYP/aug-cc-pVDZ method, and (iii) the MP2/aug-cc-pVDZ level. The reference set of values for cysteine obtained at the B3LYP/6-31G** level (Table 2) was 8.022 Å, 13.871 Å, 85.652 Å·g/mol, and 2.254 Å·e, for $^{SR}CM_0^{(p)}$, $^{SR}CM_L^{(p)}$, $^{SR}CM_M^{(p)}$, and $^{SR}CM_q^{(p)}$, respectively. By using the same B3LYP/6-31G** method, the B3LYP/aug-cc-pVDZ method and the MP2/aug-cc-pVDZ method we obtained the following sets of values: 8.020, 13.871, 85.651, and 2.255; 8.447, 14.404, 89.138, and 3.593; and 8.535, 15.021, 102.906, and 2.550 values, respectively. Thus reproducibility is perfect, the change with basis set from 6-31G** to aug-cc-pVDZ is within 5%, 3.8%, 4%, and 40%; and the method from B3LYP to MP2 (aug-cc-pVDZ basis set) is within 1%, 3.7%, 15%, and 30%. Thus, the pure and labeled chirality measures perform well with the basis and method change, the mass measure seems to be oversensitive to the change of method from DFT to MP2, whereas variation of the charge measure is unacceptable. However, it is quite clear why in the latter case it is so. For the charge chirality measure, the Mulliken charges were used, whereas it is known that they are not convergent as the basis set is increased and perform quite erratically with the change of the computational method. Thus, the defective behavior of the charge chirality measure is a consequence of the defective performance of the Mulliken charges, and this fact should be taken into account in further use of this kind of measure.

It is important that the presented methodology is naturally applicable to systems with two chiral centers. Indeed, observing the chirality measures found for all four threonine enantiomers: 2*R*,3*S*-; 2*R*,3*R*-; 2*S*,3*R*-; and 2*S*,3*S*- (Tables 2 and 16a−d). The measures were calculated at the B3LYP/6-31G** level and averaged over entire populations of conformers. As one would expect, the measures for the 2*R*,3*S*- and 2*S*,3*R*- enantiomers are identical (within the levels of computational error). For the 2*R*,3*R*- and 2*S*,3*S*- enantiomers, they are identical as well. Indeed, the measures are based on similarity between enantiomers and are identical for a molecule and its mirror image. However, the 2*R*,3*S*- and 2*R*,3*R*- isomers are not mutual mirror images, and thus, their chirality measures are definitely different (Table 2). Thus, the method is directly applicable to systems with two or more chirality centers to differentiate isomers that are not exact mirror images, whereas the exact enantiomers are placed into the class of the same measure. Potential differentiation of these could be performed by an additional arbitrary rule.

^{SR}CMs vs other Chirality Measures of α-Amino Acids. Before this study, Boon et al. studied Molecular Quantum Similarity chirality measures for selected α-amino acids (alanine, asparagine, cysteine, leucine, serine, and valine).[123−126] They derived their measures from Carbó's molecular quantum similarity theory based on the Euclidean distance between the electron densities of the two molecules under comparison. In the case of a pair of enantiomers, the Carbó similarity index becomes a chirality measure.[127,128] The Boon et al. measure for conformers had taken into account the Boltzmann distribution at given temperature, and the best overlap of the electron densities were obtained by using the

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

Table 2. Different Types of Population Averaged Chirality Measures Obtained by Using the CHIMEA Program for 19 -Amino Acids (Glycine and Cystine Omitted) Found at the B3LYP/6-31G Level[a]**

			B3LYP/6-31G**				crystal structure			Dryzun-Avnir [126]	Boon et al.[121]			
amino acid		no. conf	$^{SR}CM_0^{(p)}$	$^{SR}CM_L^{(p)}$	$^{SR}C_M^{(p)}$	$^{SR}CM_q^{(p)}$	$^{SR}CM_0^{(c)}$	$^{SR}CM_L^{(c)}$	$^{SR}CM_M^{(c)}$	S_{chiral}	BB	TGSA	BB_D	$TSGA_D$
alanine	Ala	9	6.814	11.048	47.413	2.403	3.96	11.28	59.49	9.38	0.3908	0.6930	0.4633	0.5121
arginine	Arg	520	24.772	41.349	213.327	10.670	na	na	na	2.78	nd	nd	nd	nd
aspargine	Asn	23	10.495	20.312	105.414	5.153	13.90	20.90	108.00	9.68	nd	nd	nd	nd
aspartic acid	Asp	37	8.768	15.957	100.335	4.493	10.88	15.86	95.30	8.29	0.2408	0.4445	0.3178	0.4206
cysteine	Cys	50	8.022	13.871	85.652	2.254	7.73	15.05	63.17	9.07	0.1072	0.1843	0.3953	0.4468
glutamine	Gln	72	14.738	24.992	139.493	6.152	8.67	20.80	130.87	6.99	nd	nd	nd	nd
glutamic acid	Glu	143	13.541	20.231	119.785	5.185	13.76	22.86	137.28	14.13	nd	nd	nd	nd
histidine	His	38	11.946	25.524	155.071	6.132	12.68	18.13	105.18	7.68	nd	nd	nd	nd
isoleucine	Ile	59	15.518	29.092	141.409	5.968	13.53	27.65	124.15	7.73	nd	nd	nd	nd
leucine	Leu	53	13.544	19.042	88.015	3.456	12.87	20.68	104.48	6.49	0.3075	0.3419	0.3407	0.4304
lysine	Lys	391	15.937	24.140	116.114	4.766	na	na	na	3.15	nd	nd	nd	nd
methionine	Met	113	14.359	22.788	112.571	4.609	na	na	na	4.17	nd	nd	nd	nd
phenylalanine	Phe	20	11.395	17.756	133.657	5.022	8.20	12.45	102.05	4.83	nd	nd	nd	nd
proline	Pro	3	7.476	16.690	77.601	3.432	6.80	18.83	103.30	14.01	nd	nd	nd	nd
serine	Ser	30	6.964	11.786	59.579	2.738	5.62	15.22	79.99	17.25	0.3050	0.5715	0.3521	0.4388
2S,3R-threonine	Thr	29	10.367	17.099	85.238	4.127	10.33	14.13	77.27	7.65	nd	nd	nd	nd
2R,3S- threonine		29	10.348	17.103	86.629	4.197								
2S,3S- threonine		32	9.152	16.188	67.789	3.187								
2*R*,3*R*- threonine		32	9.098	16.286	68.283	3.203								
tryptophan	Trp	37	21.729	34.757	205.845	7.229	20.10	29.50	239.40	5.46	nd	nd	nd	nd
tyrosine	Tyr	38	12.934	20.327	150.926	6.368	11.68	25.60	178.40	4.03	nd	nd	nd	nd
valine	Val	19	11.267	16.206	90.905	4.037	8.76	14.44	75.14	7.10	0.3219	0.4485	0.3242	0.5889

[a]The units of the $^{SR}CM_0^{(p)}$, $^{SR}CM_L^{(p)}$, $^{SR}CM_M^{(p)}$, and $^{SR}CM_q^{(p)}$ measures are the following: angstroms, angstroms, angstrom gram per mole, and angstrom electron charge, respectively. The superscript (p) denotes the population weighted measure, while (c) denotes the average over all studied crystal data. Populations (p) were calculated from Gibbs free energy differences (kilocalories per mole, 298.15 K, 1 atm). BB and TGSA stand for backbone and topo-geometrical superposition algorithm of search of maximal of electron density between enantiomers, and subscript D denotes density difference function. na stands for not available due to the lack of reliable crystallographic data; nd stands for not determined.

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

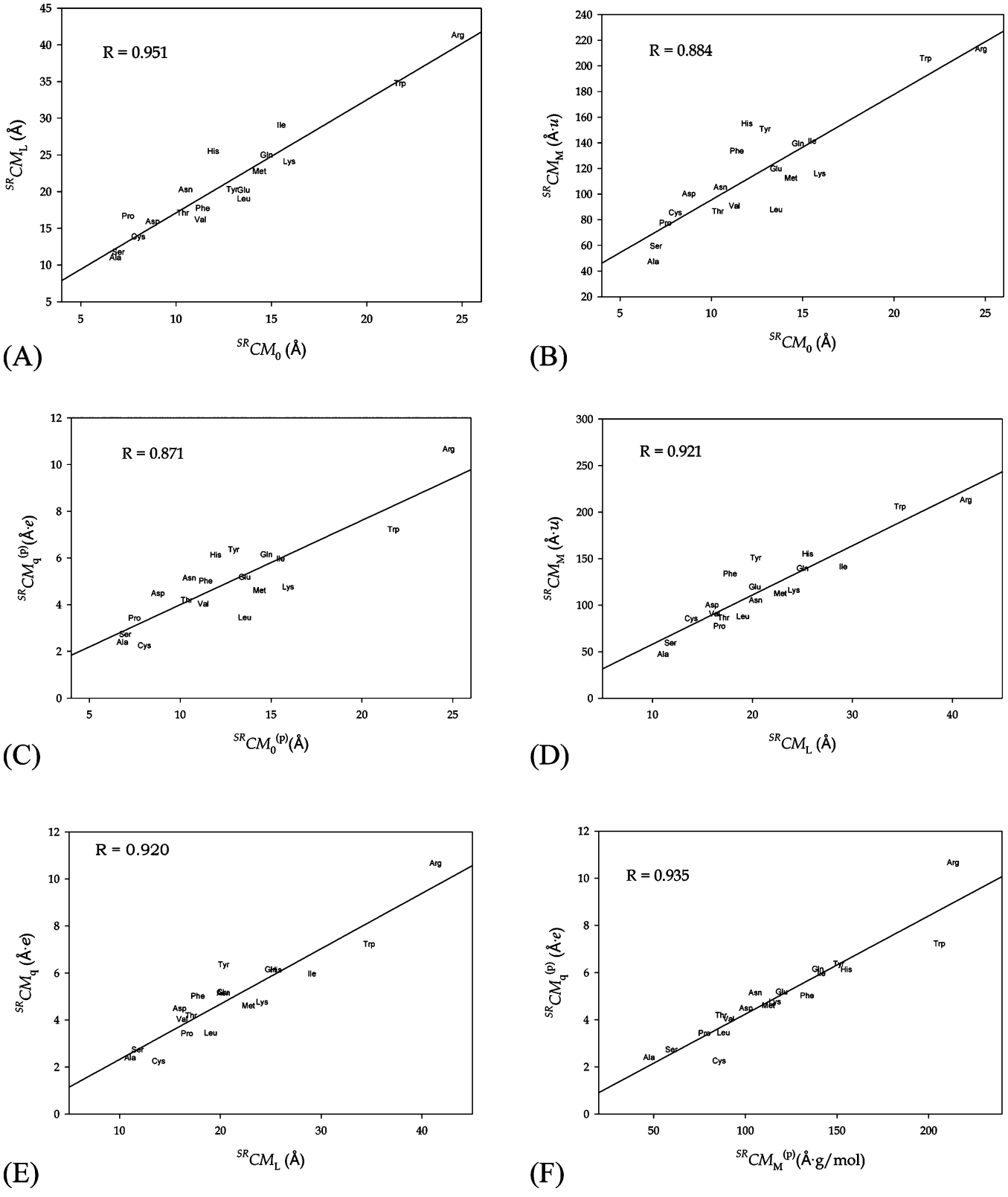


Figure 3. Mutual linear correlations between chirality measures calculated for amino acids in the gas phase (averaged over population estimated for standard condition at the B3LYP/6-31G** level). $^{SR}CM_0^{(p)}$, $^{SR}CM_L^{(p)}$, $^{SR}CM_M^{(p)}$, and $^{SR}CM_q^{(p)}$ denote the pure, labeled, mass, and charge chirality measure.

backbone (BB) and topo-geometrical superposition algorithm (TGSA) (which is based on comparisons of atom types and interatomic distances, allowing for handling large molecular sets within affordable computational costs).[129] Thus, a similarity of enantiomer is a common base of our study and that of Boon et al., and also, their method of comparison seems to be similar to our labeled measure. Unfortunately, comparison between our four chirality measures with their chirality measures obtained by

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

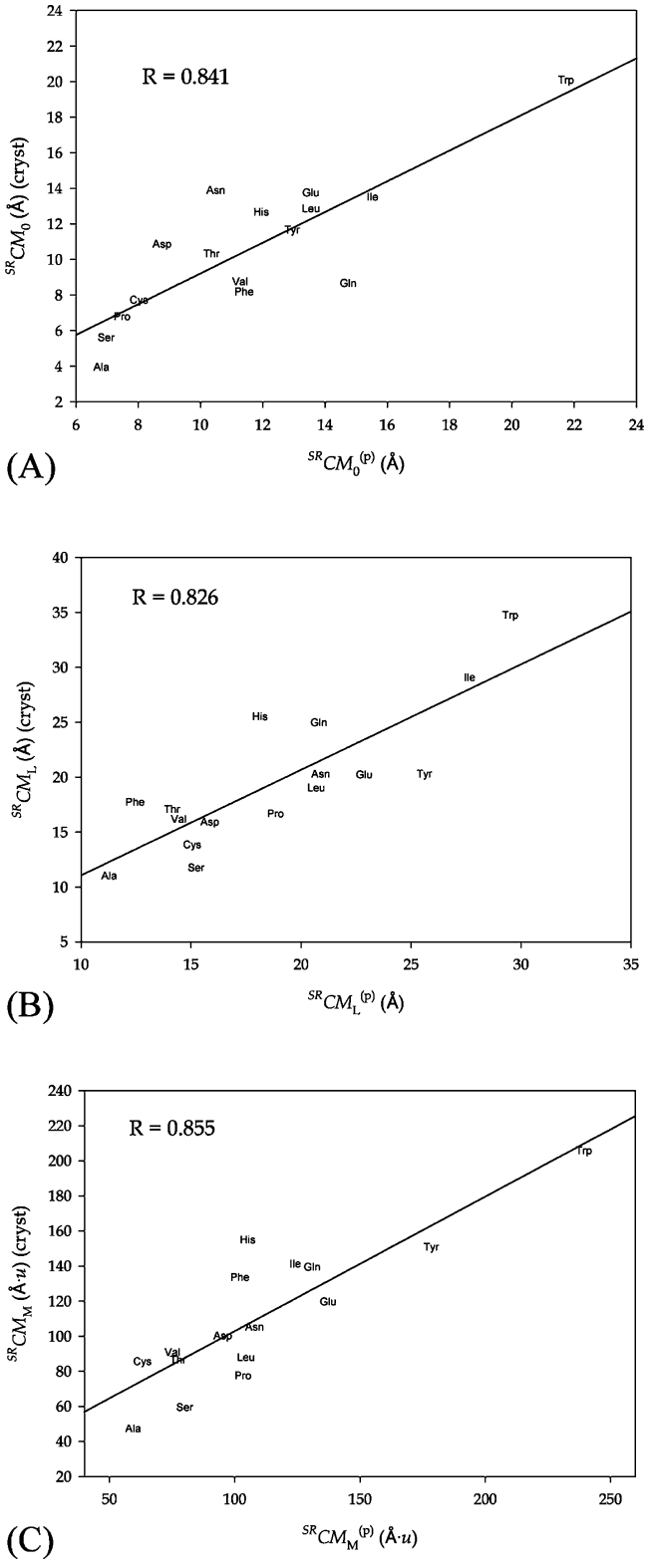


Figure 4. Linear correlations between chirality measures calculated for zwitterionic forms of amino acids in crystals (averaged over all available data for given crystal) and population averaged chirality measures calculated for molecules in the gas phase at the B3LYP/6-31G** level. $^{SR}CM_0$, $^{SR}CM_L$, and $^{SR}CM_M$ denote the pure, labeled, and mass chirality measure.

the two enantiomer comparison routines applied to both electron densities and density difference function (Table 2) in most cases yields, more or less, a chaotic spread of points in the plots. One of the possible causes of such a spread is the low number of compared points, being equal to six. For such a small set, even one scattered point destroys a possible correlation. This may be the case for most of the plots between our data and the measures of Boon et al. The other probable cause for the lack of correlations is the fact that Boon et al. measures practically do not intercorrelate between themselves. Again, this may be a consequence of the small number of compared points, as well as an instability of the minimization routines. However, at the moment, explanation as to whether there is a correlation between our measures and the measures of Boon et al. goes beyond the limits of this study.

Very recently, Dryzun and Avnir have generalized the Continuous Chirality Measures concept to vectors, matrices, operators, and functions, extending the possibility for calculation of the chirality content to any mathematical description of a chiral system by vectors, matrices, operators, and functions.[130] They applied the newly developed methodology to amino acids and listed chirality measures for all α-amino acids apart from cystine (Table 2). However, the strongest correlation between Dryzun and Avnir chirality measures and our chirality measures is for the pure measure, and it still is weak (Figure 5). There may be several causes of

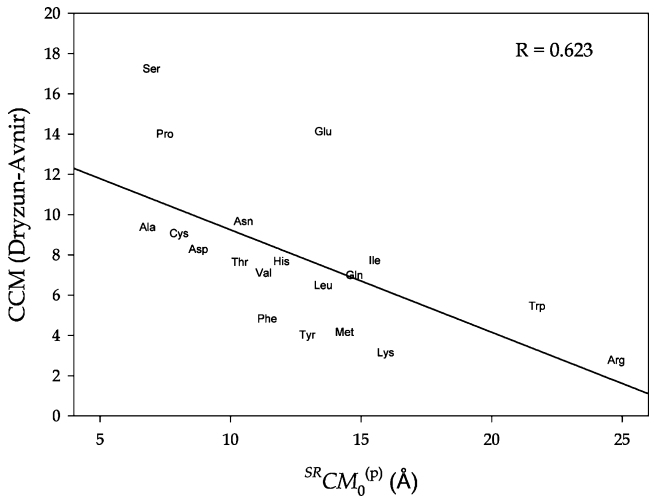


Figure 5. Weak linear correlation between Dryzun and Avnir CCM chirality measures for α-amino acids and pure $^{SR}CM_0^{(p)}$ chirality measure (averaged over population of amino acids in the gas phase) calculated at the B3LYP/6-31G** level.

these weak correlations. The main one is that the measures were obtained by different approaches: Dryzun–Avnir values are found from a distance of a chiral molecule from a symmetrical achiral set, whereas our data are connected to dissimilarity between enantiomers. If intercorrelations even between our chirality measures derived from the same ground are not perfect, it is not very surprising that the correlations between the measures derived from different assumptions are weaker (Figure 5). However, there may be another reason for weak correlation between these measures. The Dryzun–Avnir values seem to not include any averaging over the population of conformers, and their values would be significantly changed if such an averaging were performed. Yet, such a task would require performing of exhausting calculations for a similar number of conformers to that used in this study.

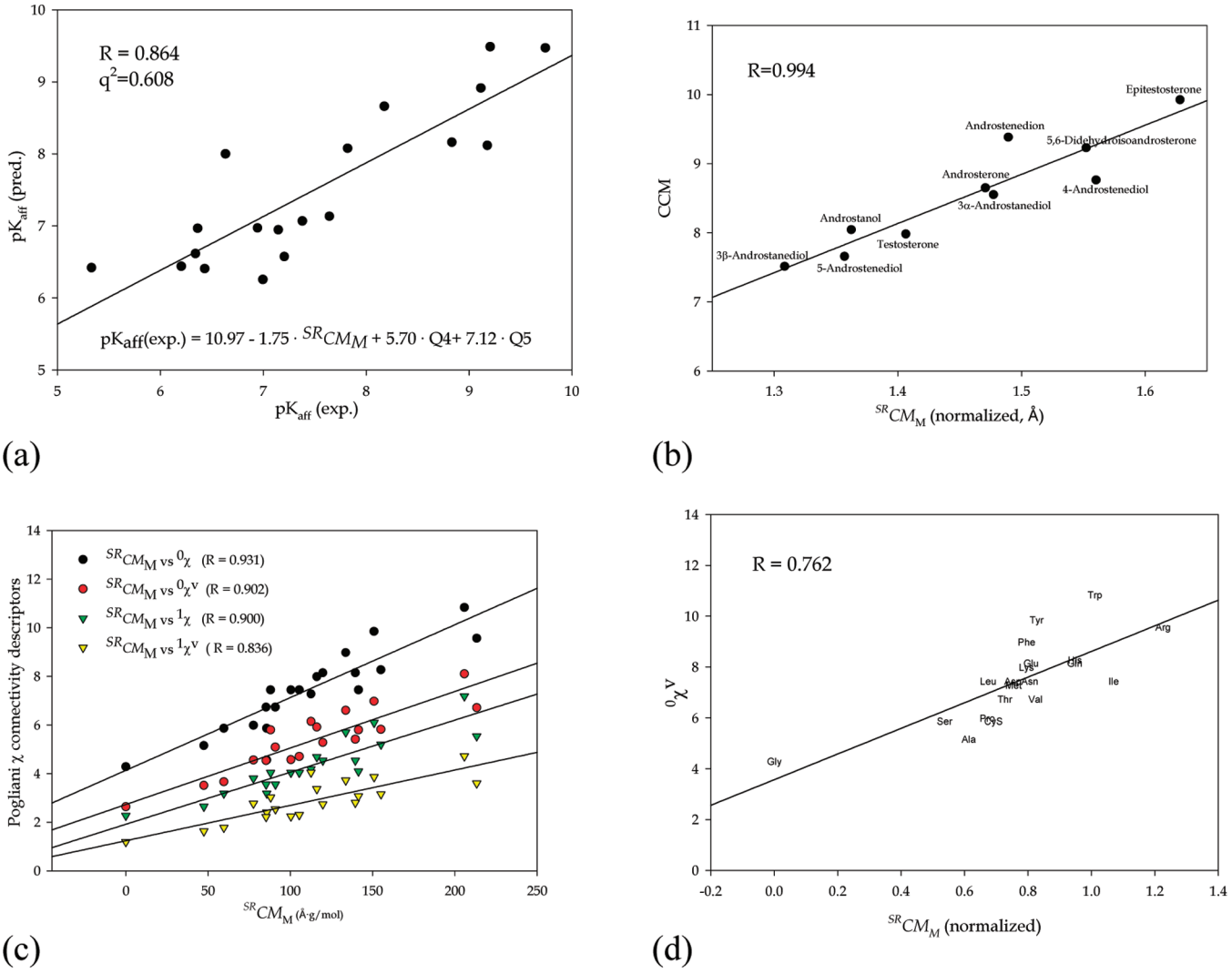


Figure 6. (a) QSAR model for the binding affinity (pK_{aff}) of the steroidal ligands for the sex-hormone binding globulin:[131] Q4 and Q5 descriptors the partial charges of the C4 and C5 atoms of steroidal skeleton determined by the CHELPG method and the B3LYP/6-31G calculations. (b) Linear regression between the CCM[130] and normalized $^{SR}CM_M$ chirality measures for various androgens (Scheme 1). (c) Linear regression between the un-normalized $^{SR}CM_M$ chirality measure of the α-amino acids and connectivity indices.[139] (d) Weak linear regression between the normalized $^{SR}CM_M$ chirality measure of the α-amino acids and the $^{0}\chi^{v}$ connectivity index.[139]

Perspectives for ^{SR}CMs as QSAR Descriptors. It is acknowledged that the drug action depends mainly on protein–ligand interactions where a geometrical fit is of great importance. Thus, the shape and other spatial properties (including chirality, skewedness, or asymmetry) of an active molecule may have great impact on its activity. So far, among various QSAR descriptors to account for these properties, only a few are directly related to chirality and their use in QSAR equations.[33−35]

A sample of possible QSAR applications of the ^{SR}CM descriptors can be given by preliminary QSAR equations obtained for a set of sex-hormone binding globulin (SHBG) ligands. An example model (Figure 6a) was built for the benchmark Cramer data set containing 21 SHBG ligands[131] (two compounds were removed from the set: etiocholanone because of its structural unconformity and androsterone because of computational instability). The models were prepared so far without external validation and statistically are only acceptable ($R = 0.864$, $q^2 = 0.608$ where q stands for leave-one-out cross-validation correlation coefficient). Indeed, the accepted limits of statistical significance, $q^2 = 0.5$[132] is only slightly lower than the obtained values. Nevertheless, these introductory results show that ^{SR}CMs can be successfully introduced into QSAR equations. The study is in progress and includes a much larger set of molecules,[133] molecular descriptors, and external validation. At the moment, we are working upon including ^{SR}CMs into QSAR models for sex-hormone binding globulin ligands and the androgen receptor, since the great differences in binding affinity of structurally related ligands[134] cannot be well explained based only on distinct hydrogen bond patterns.[135]

It is nearly certain that different atomic properties are crucial in different QSAR problems. The ^{SR}CM chirality measures are constructed to easily convolute a desired atomic property with the molecular chirality measure. However, there are important issues that must be considered before ^{SR}CMs can be consciously used in QSAR analyses. First, are there any connections between ^{SR}CMs and the other QSAR descriptors? Second, which of these pure, labeled, mass, or charge ^{SR}CM is the most

dx.doi.org/10.1021/ci300057h | J. Chem. Inf. Model. 2012, 52, 1462−1479

appropriate to model particular QSAR problems? And third, is the normalization factor important to QSAR analyses or not?

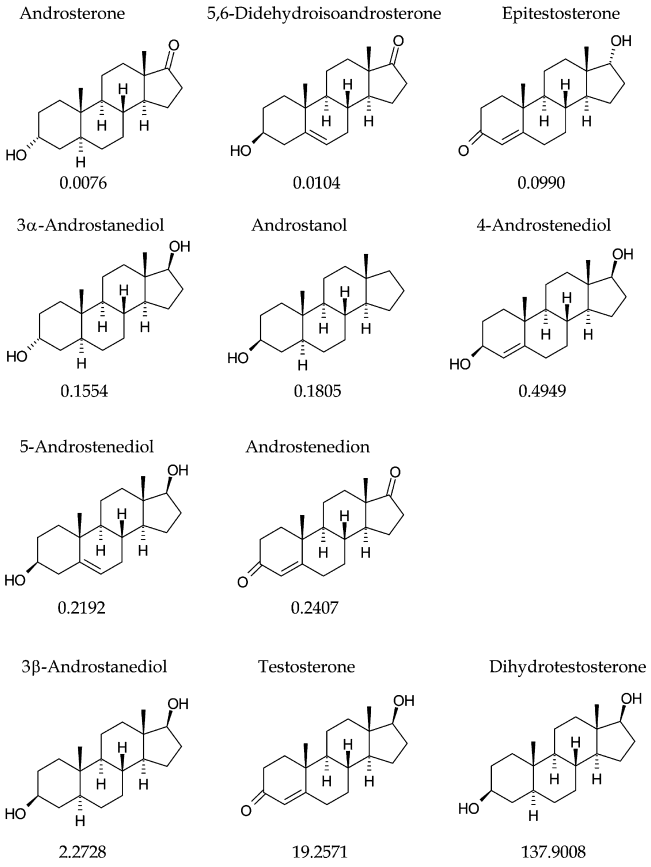
To initially address these issues, let us first demonstrate perfect correlation (R = 0.994) between the Avnir group CCM chirality measure[14] and the normalized $^{SR}CM_M$ measure (Figure 6b), obtained for 10 sex-hormone binding globulin ligands (Scheme 1). Interestingly, the un-normalized $^{SR}CM_M$ performs equally well. The analogous correlation with the $^{SR}CM_q$ measure is weaker but definitely statistically significant (R = 0.928), whereas the ones with the other ^{SR}CM descriptors are insignificant. We think that the specific structure of steroids, being relatively flat and rigid molecules, may be the reason why the CCM and $^{SR}CM_M$ descriptors correlate. Especially for such kinds of molecules, distance to an abstract achiral planar object (CCM) can be linearly related to a distance between two, somehow planar, enantiomers ($^{SR}CM_M$). Yet, the reason why the $^{SR}CM_M$ descriptor behaves this way, while the other ^{SR}CM descriptors do not follow the same tendency, remains unknown as yet. Although the CCM descriptor is still not widely used in QSAR/QSPR analyses, it was shown to have great potential.[136−138] Thus, although for α-amino acids the correlation between CCM and $^{SR}CM_M$ is only weak (Figure 5), it is quite probable that in some cases where the CCM performs well, the ^{SR}CMs may also perform well.

Unexpectedly, we found quite significant linear correlation between the connectivity indices published by Pogliani for α-amino acids,[139] and the un-normalized $^{SR}CM_M$ (R = 0.931, Figure 6c). The analogous correlation with the Randić type of variable connectivity index[140] for α-amino acids[141] is also significant but a bit weaker. This is surprising because the connectivity indices express first and foremost the molecular branching[142] and the molecular accessibility,[143] the properties which are hardly connected to molecular chirality. Moreover, significant linear regressions with connectivity indices implicate significant linear regressions with a variety of chirality-unrelated physical properties modeled for amino acids by these very indices.[139,141,144−150] However, it turned out that the correlations between the same connectivity indices and the normalized $^{SR}CM_M$ are weak, if at all significant ($R \approx 0.7$, Figure 6d).

It is of key importance to understand why, for amino acids, there is a good correlation with the un-normalized $^{SR}CM_M$ but there is not a significant correlation with the normalized $^{SR}CM_M$. The un-normalized $^{SR}CM_M$ contains factors proportional to masses of atoms constituting the chiral molecule (eq 2a), and thus, it is proportional to the size of the molecule, its spatial volume, and related properties which are well modeled by connectivity indices.[139,141,144−150] Information on molecular size (mass) is released from the descriptor by normalization of the $^{SR}CM_M$, and in consequence, the correlations with molecular size and topology become insignificant (Figure 6d). Why, however, for the 10 sex-hormone binding globulin ligands are the correlations of CCM and $^{SR}CM_M$ equally good regardless of whether the $^{SR}CM_M$ were normalized or not? This is because these molecules (Scheme 1) are of similar mass, and the mass factor becomes a kind of a constant ratio between the normalized and un-normalized descriptor.

Thus, if the un-normalized descriptors share so much information with other descriptors connected to chirality-unrelated molecular properties, should normalization not be the necessary option for use of ^{SR}CMs in QSAR? The answer to this question is not as straightforward as one might think. First, it may be profitable for QSAR analyses that one descriptor combines two properties in one value because it could decrease the number of variables used. Second, other ^{SR}CMs are normalized differently than the $^{SR}CM_M$. For example, in case of the $^{SR}CM_q$ descriptor, the charges are of both signs and in normalization they are taken as absolute values. Their sum expresses, in kind, the overall polarity of the molecule, which has not yet been tested for its connections with the other QSAR descriptors. Third, the $^{SR}CM_M$ normalized values approach zero as, for example, the aliphatic chain in series of homologues increases to infinity, whereas the un-normalized values remain roughly constant. The former property may be important in some applications, while the latter in other ones. Therefore, much more studies are needed to fully understand the role of chirality measures as QSAR descriptors.

Scheme 1. Various Androgens of Similar Structure with Different Affinity for the Androgen Receptor[a]



[a]The affinity is described as the relative binding affinity.[131]

CONCLUSIONS

The concept of chirality measures was developed through the concept of Property Space and similarity between enantiomers. The Property Space $R^3 \times P^k$ is a Cartesian product of the Euclidean R^3 space, in which a molecule is embedded, and the proper Property Space P^k of features that can be assigned to individual atoms in a molecule, such as type of atom, mass or charge. To calculate the measures, the CHIMEA computer program was written. The inputs for the program can be either from the protein databank or in Gaussian log format and four basic chirality measures: pure, labeled, mass, and charge, are given in the output. The program minimizes the distance

between enantiomers in the assumed Property Space. The pure, labeled, mass, and charge chirality measures differ in their weight of modifying the expression determining distance between enantiomers in the Property Space.

The pure, labeled, mass, and charge chirality measures were calculated for all α-amino acid zwitterion structures for which the molecular structures in crystals were solved with precision, enabling determination of all positions of H-atoms. The calculations were also performed for the amino acid conformers, except for glycine and cystine, calculated at the B3LYP/6-31G** level. As many as 1684 different stable conformers of all studied amino acids were considered and populations in standard conditions were calculated for them, based on Gibbs free energies calculated at the same level of theory.

The different kinds of measures were tested for mutual correlations for both crystal zwitterionic averaged over structures in all available crystals and neutral amino acids measures averaged over the population at 298 K. The measures for α-amino acid zwitterions in the crystal phases intercorrelate, but not strongly: R ranges from 0.78 to 0.86. The factors that influences the quality of such correlations are connected with the fact that they are independent measures and that the crystals considered differ in both the structure determination quality parameters, and combinatorial differences of finding the global minimum in search of the optimal measure. However, mutual correlations between the measures found for amino acids in the gas phase (averaged over conformer populations) are much better: R ranges from 0.87 to 0.95. It was found that the correlations between the labeled measure and mass and charge, as well as between mass and charge, perform quite well, and for them the correlation coefficient exceeds 0.920. We found also that there are weak, yet significant, correlations between chirality measures of amino acids in the gas phase and analogous measures for amino acid zwitterions in the crystal phase, strongly suggesting that the chirality measures obtained based on our calculations can be valuable parameters in QSAR analyses for molecules operating in different phases of condensation and ordering.

To better determine performance of the measures and the CHIMEA program, we checked the reproducibility of the chirality values by repeating the calculations for the set of 50 cysteine conformers calculated at the B3LYP/6-31G** level was perfect. We also recalculated the set of 50 cysteine conformers at the B3LYP/aug-cc-pVDZ and MP2/aug-cc-pVDZ levels, showing that the pure and labeled chirality measures perform well with the basis and calculation method change, the mass measure seems to be oversensitive to change of the method from DFT to MP2, whereas variation of the charge measure is unacceptable because the Mulliken charges, used in the calculations, perform quite erratically with the change of the computational method. Thus, proper selection of the charge determination method, convergent as the computation level is increased, is recommended in future studies.

Finally, we compared the set of chirality measures obtained in this study for α-amino acids with chirality measures obtained by Boon et al. and Dryzun and Avnir. We found no satisfactory correlations between these data, which, however, may be a consequence of different methodologies used in these studies. Also, we discussed the perspectives for ^{SR}CMs as QSAR descriptors. We presented an example QSAR model for a set of steroidal SHBG ligands showing that ^{SR}CMs can be successfully introduced into QSAR equations. The possible relations between ^{SR}CMs and known QSAR descriptors were also discussed.

■ ASSOCIATED CONTENT

Supporting Information

Tables 1SI−19SI as mentioned in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: janek@il.waw.pl.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partially supported by a grant from the Ministry of Science and Higher Education in Poland for statutory activity of the Industrial Chemistry Research Institute in 2010 and 2011. The computational Grants G19-2 and G19-4 from the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the Warsaw University are gratefully acknowledged. We are very grateful to anonymous reviewers for their constructive criticism and suggestions to improve the manuscript.

■ REFERENCES

(1) Pasteur, L. Recherches sur les relations qui peuvent exister entre la forme cristalline, la composition chimique et le sens de la polarisation rotatoire. *Annal. chim. phys., 3e sér.* **1848**, *24*, 442−459 ; "*Je montre, en effet, que l'hémiédrie est liée avec le sens de la polarisation rotatoire. Or, ce dernier phénomène étant moléculaire et accusant une dissymétrie dans les molécules, l'hémiédrie, à son tour, se trouve donc en étroite connexion avec la dissymétrie des derniers éléments qui composent le cristal*".

(2) Flack, H. D. Louis Pasteur's discovery of molecular chirality and spontaneous resolution in 1848, together with a complete review of his crystallographic and chemical work. *Acta Crystallogr.* **2009**, *A65*, 371−389.

(3) Van't Hoff, J. H. Voorstel tot uitbreiding der structuur formules in de ruimte, Sur les formules de structure dans l'espace. *Arch. Neerl. Sci. Exact. Nat.* **1874**, *9*, 445−454.

(4) Le Bel, J. A. Sur les relations qui existent entre les formules atomiques des corps organiques et le pouvoir rotatoire de leurs dissolutions. *Bull. Soc. Chim. Fr.* **1874**, *22*, 337−347.

(5) Kelvin, W. T. The molecular tactics of a crystal. *J. Oxford Univ. Jr. Sci. Club* **1894**, *18*, 3−57.

(6) Bentley, R. Chiral: A Confusing Etymology. *Chirality* **2010**, *22*, 1−2.

(7) Basic terminology of stereochemistry. *IUPAC Recommendations 1996*; PAC, 1996; Vol. *68*, pp 2193 and 2203, blue book, p 479.

(8) Petitjean, M. Chirality and Symmetry Measures: A Transdisciplinary Review. *Entropy* **2003**, *5*, 271−312.

(9) Buda, A. B.; Auf der Heyde, T.; Mislow, K. On Quantifying Chirality. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 989.

(10) Kitaigorodskii, A. I. *Organic Chemical Cristallography*; Consultans Bureau Ets. Inc., New York, 1961.

(11) Gilat, G.; Schulman, L. S. Chiral Interaction, Magnitude of Effects and Application to Natural Selection of L-Enantiomer. *Chem. Phys. Lett.* **1985**, *121*, 13−16.

(12) Zabrodsky, H.; Peleg, S.; Avnir, D. Continuous Symmetry Measures. *J. Am. Chem. Soc.* **1992**, *114*, 7843−7851.

(13) Zabrodsky, H.; Peleg, S.; Avnir, D. Continuous Symmetry Measures. 2. Symmetry Groups and the Tetrahedron. *J. Am. Chem. Soc.* **1993**, *115*, 8278−8289.

(14) Zayit, A.; Pinsky, M.; Elgavi, H.; Dryzun, C.; Avnir, D. A web site for calculating the degree of chirality. *Chirality* **2011**, *23*, 17−23, http://www.csm.huji.ac.il/.
(15) Rassat, A. Un critére de classement des systèmes chireau de points à partir de la distance au sence de Hausdorff. *Compt. Rend. Acad. Sci. Paris (Sér. II)* **1984**, *299*, 53−55.
(16) Buda, A. B.; Mislow, K. A Hausdorff chirality measure. *J. Am. Chem. Soc.* **1992**, *114*, 6006−6012.
(17) Mezey, P. G. The proof of the metric properties of a fuzzy chirality measure of molecular electron density clouds. *J. Mol. Struct. (Theochem)* **1998**, *455*, 183−190.
(18) Randić, M. Graph Theoretical Descriptors of Two-Dimensional Chirality with Possible Extension to Three-Dimensional Chirality. *J. Chem. Inf. Comp. Sci.* **2001**, *41*, 639−649.
(19) Natarajan, R.; Basak, S. C.; Neumann, T. S. Novel Approach for the Numerical Characterization of Molecular Chirality. *J. Chem. Inf. Model.* **2007**, *47*, 771−775.
(20) Golbraikh, A.; Bonchev, D.; Tropsha, A. Novel Chirality Descriptors Derived from Molecular Topology. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 147−158.
(21) Aires-de-Sousa, J.; Gasteiger, J.; Gutman, I.; Vidović, D. Chirality Codes and Molecular Structure. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 831−836.
(22) Petitjean, M. About second kind continuous chirality measures.1. Planar sets. *J. Math. Chem.* **1997**, *22*, 185−201.
(23) Petitjean, M. On the Root Mean Square Quantitative Chirality and Quantitative Symmetry Measures. *J. Math. Phys* **1999**, *40*, 4587−4595.
(24) Petitjean, M. Calcul de chiralité quantitative par la mkthode des moindres carrés. *C. R. Acad. Sci. Paris, t. 2, Sér. II* **1999**, 25−28.
(25) Petitjean, M. Chiralité quantitative: le modèle des moindres carrés pondérés. *Compt. Rend. Acad. Sci. Paris, Sér. IIc* **2001**, *4*, 331−333.
(26) Petitjean, M. Chiral Mixtures. *J. Math. Phys.* **2002**, *43*, 4147−4157.
(27) Petitjean, M. À propos de la référence achirale. *Compt. Rend. Chim* **2006**, *9*, 1249−1251.
(28) Petitjean, M. Chirality in Metric Spaces Symmetry: Culture and Science. 2010, *21*, 27−36.
(29) Avnir, D.; Hel-Or, H.; Mezey, P. *Symmetry and Chirality: Continuous Measures in The Encyclopedia of Computational Chemistry*; Schleyer, P. V., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol *4*, pp 2890−2901; http://cs.haifa.ac.il/hagit/papers/CCencyclopedia9.
(30) Casanova i Casas, D. *Mesures de forma i simetria en química: algorismes i aplicacions*. Tesis doctoral dirigida por Reverter, S. Á., Cahner, P. A.; Universitat de Barcelona, 2006.
(31) Natarajan, R.; Basak, S. C. Numerical characterization of molecular chirality of organic compounds. *Cur. Comput.-Aided Drug Design* **2009**, *5*, 13−22.
(32) Katritzky, A. R.; Kuanar, M.; Slavov, S.; Hall, C. D.; Karelson, M.; Kahn, I.; Dobchev, D. A. Quantitative Correlation of Physical and Chemical Properties with Chemical Structure: Utility for Prediction. *Chem. Rev.* **2010**, *110*, 5714−89.
(33) Natarajan, R.; Basak, S. C. Numerical Descriptors for the Characterization of Chiral Compounds and their Applications in Modeling Biological and Toxicological Activities. *Cur. Top. Med. Chem.* **2011**, *11*, 771−787.
(34) Basak, S. C. Role of Mathematical Chemodescriptors and Proteomics-Based Biodescriptors in Drug Discovery. *Drug Dev. Res.* **2011**, *72*, 225−233.
(35) Crippen, G. M. Chirality Descriptors in QSAR. *Cur. Comput.-Aided Drug Design* **2008**, *4*, 259−264.
(36) Jamróz, M. H. *CHIMEA - program calculating discrete chirality measures of molecules*; Warsaw, 2010.
(37) Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr.* **2002**, *B58*, 380−388.
(38) van de Streek, J. Searching the Cambridge Structural Database for the 'best' representative of each unique polymorph. *Acta Crystallogr.* **2006**, *B62*, 567−579.
(39) Orpen, A. G. Applications of the Cambridge Structural Database to molecular inorganic chemistry. *Acta Crystallogr.* **2002**, *B58*, 398−406.
(40) Allen, F. H.; Motherwell, W. D. S. Applications of the Cambridge Structural Database in organic chemistry and crystal chemistry. *Acta Crystallogr.* **2002**, *B58*, 407−422.
(41) Taylor, R. Life Science applications of the Cambridge Structural Database. *Acta Crystallogr.* **2002**, *D58*, 879−888.
(42) *Spartan '08*; Wavefunction, Inc.: Irvine, CA; www.wavefun.com.
(43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision B.01*; Gaussian, Inc., Wallingford CT, 2009.
(44) Cahn, R. S.; Ingold, C. K.; Prelog, V. Specification of Molecular Chirality. *Angew. Chem. Int. Ed. Eng.* **1966**, *5*, 385−415 (errata: 1966, *5*, 511).
(45) Prelog, V.; Helmchen, G. Basic Principles of the CIP+Szstem and Proposals for a Revision. *Angew. Chem. Int. Ed. Eng.* **1982**, *21*, 567−583.
(46) Ostrowski, S.; Jamróz, M. H.; Rode, J. E.; Dobrowolski, J. Cz. On Stability, Chirality Measures, and Theoretical VCD Spectra of the Chiral $C_{58}X_2$ Fullerenes (X = N, B). *J. Phys. Chem.* **2012**, *116*, 631−643.
(47) Wilson, C. C.; Myles, D.; Ghosh, M.; Johnson, L. N.; Wang, W. Neutron diffraction investigations of L- and D-alanine at different temperatures: the search for structural evidence for parity violation. *New J. Chem.* **2005**, *29*, 1318−1322.
(48) Subha Nandhini, M.; Krishnakumar, R. V.; Natarajan, S. DL-Alanine. *Acta Crystallogr. C.* **2001**, *57*, 614−615.
(49) Simpson, H. J.; Marsh, R. E. The crystal structure of L-alanine. *Acta Crystallogr.* **1966**, *20*, 550−555.
(50) Dunitz, J. D.; Ryan, R. R. Refinement of the L-alanine crystal structure. *Acta Crystallogr.* **1966**, *21*, 617−618.
(51) Destro, R.; Marsh, R. E.; Bianchi, R. A low-temperature (23 K) study of L- alanine. *J. Phys. Chem.* **1988**, *92*, 966−973.
(52) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. I. Crystal and molecular structure of the amino acid L-alanine. *J. Am. Chem. Soc.* **1972**, *94*, 2657−2660.
(53) Wang, W.; Gong, Y.; Li, C. Biochirality and parity violating energy difference. *Chin. Sci. Bull.* **2002**, *47*, 603−608.
(54) Dalhus, B.; Gorbitz, C. H. Molecular aggregation in selected crystalline 1:1 complexes of hydrophobic D- and L-amino acids. III. The L-leucine and L-valine series. *Acta Crystallogr. C* **1999**, *55*, 1547−1555.
(55) Dalhus, B.; Gorbitz, C. H. Crystal Structures of Hydrophobic Amino Acids l. Redeterminations of L-Methionine and L-Valine at 120 K. *Acta Chem. Scand.* **1996**, *50*, 544−548.
(56) Mallikarjunan, M.; Rao, S. T. Crystal structure of DL-valine. *Acta Crystallogr. B.* **1969**, *25*, 296−303.
(57) Dalhus, B.; Gorbitz, C. H. Triclinic Form of dl-Valine. *Acta Crystallogr. C.* **1996**, *52*, 1759−1761.
(58) Flaig, R.; Koritsánszky, T.; Dittrich, B.; Wagner, A.; Luger, P. Intra- and Intermolecular Topological Properties of Amino Acids: A

Comparative Study of Experimental and Theoretical Results. *J. Am. Chem. Soc.* **2002**, *124*, 3407−3417.
(59) Dalhus, B.; Gorbitz, C. H. Structural relationships in crystals accommodating different stereoisomers of 2-amino-3-methylpentanoic acid. *Acta Crystallogr. B.* **2000**, *56*, 720−727.
(60) Varughese, K. I.; Srinivasan, R. Studies in molecular structure, symmetry and conformation XII: Crystal and molecular structure of D-alloisoleucine. *J. Cryst. Mol. Struct.* **1975**, *5*, 317−328.
(61) Gorbitz, C. H.; Dalhus, B. L-Isoleucine, Redetermination at 120K. *Acta Crystallogr. C.* **1996**, *52*, 1464−1466.
(62) di Blasio, B.; Pedone, C.; Sirigu, A. DL-Leucine. *Acta Crystallogr. B.* **1975**, *31*, 601−602.
(63) Coll, M.; Solans, X.; Font-Altaba, M.; Subirana, J. A. Structure of L-leucine: a redetermination. *Acta Crystallogr. C.* **1986**, *42*, 599−601.
(64) Gorbitz, C. H.; Dalhus, B. Redetermination of L-Leucine at 120K. *Acta Crystallogr. C.* **1996**, *52*, 1754−1756.
(65) Yamashita, M.; Inomata, S.; Ishikawa, K.; Kashiwagi, T.; Matsuo, H.; Sawamura, S.; Kato, M. A high-pressure polymorph of l-α-leucine. *Acta Crystallogr. E.* **2007**, *63*, o2762−o2764.
(66) Weissbuch, I.; Frolow, F.; Addadi, L.; Lahav, M.; Leiserowitz, L. Oriented crystallization as a tool for detecting ordered aggregates of water-soluble hydrophobic .alpha.-amino acids at the air-solution interface. *J. Am. Chem. Soc.* **1990**, *112*, 7718−7724.
(67) Kayushina, R. L.; Vainshtein, B. K. Rentgenografnye Opredel'enie strukturi L-prolina. *Kristallografiya* **1965**, *10*, 833−844.
(68) Myung, S.; Pink, M.; Baik, M.-H.; Clemmer, D. E. DL-Proline. *Acta Crystallogr. C.* **2005**, *61*, o506−o508.
(69) Hayashi, Y.; Matsuzawa, M.; Yamaguchi, J.; Yonehara, S.; Matsumoto, Y.; Shoji, M.; Hashizume, D.; Koshino, H. Large Non Linear Effect Observed in the Enantiomeric Excess of Proline in Solution and That in the Solid State. *Angew. Chem., Int. Ed.* **2006**, *45*, 4593−4597.
(70) Mostad, A.; Rømming, C. Crystal Structure of DL-Tyrosine. *Acta Chem. Scand.* **1973**, *27*, 401−410.
(71) Leahey, A.; Olmstead, M. M. Private Communication to CSD, 2001.
(72) Mostad, A.; Nissen, H. M.; Rømming, C. Crystal Structure of L-Tyrosine. *Acta Chem. Scand.* **1972**, *26*, 3819−3833.
(73) Frey, M. N.; Koetzle, T. F.; Lehmann, M. S.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. X. A comparison between the crystal and molecular structures of L-tyrosine and L-tyrosine hydrochloride. *J. Chem. Phys.* **1973**, *58*, 2547−2556.
(74) Luger, P.; Weber, M. DL-Cysteine at 298K. *Acta Crystallogr. C.* **1999**, *55*, 1882−1885.
(75) Gorbitz, C. H.; Dalhus, B. L-cysteine, monoclinic form, redetermination at 120 K. *Acta Crystallogr. C.* **1996**, *52*, 1756−1759.
(76) Moggach, S. A.; Clark, S. J.; Parsons, S. L-cysteine-I at 30 K. *Acta Crystallogr. E* **2005**, *61*, o2739−o2742.
(77) Moggach, S. A.; Allan, D. R.; Clark, S. J.; Gutmann, M. J.; Parsons, S.; Pulham, C. R.; Sawyer, L. High-pressure polymorphism in L-cysteine: the crystal structures of L- cysteine-III and L-cysteine-IV. *Acta Crystallogr. B.* **2006**, *62*, 296−309.
(78) Oughton, B. M.; Harrison, P. M. The crystal structure of hexagonal L-cystine. *Acta Crystallogr.* **1959**, *12*, 396−404.
(79) Chaney, M. O.; Steinrauf, L. K. The crystal and molecular structure of tetragonal L-cystine. *Acta Crystallogr. B.* **1974**, *30*, 711−716.
(80) Dahaoui, S.; Pichon-Pesme, V.; Howard, J. A. K.; Lecomte, C. CCD Charge Density Study on Crystals with Large Unit Cell Parameters: The Case of Hexagonal l-Cystine. *J. Phys. Chem. A* **1999**, *103*, 6240−6250.
(81) Moggach, S. A.; Allan, D. R.; Parsons, S; Sawyer, L.; Warren, J. E. The effect of pressure on the crystal structure of hexagonal L-cystine. *J. Synchrotron Radiat.* **2005**, *12*, 598−607.
(82) Shoemaker, D. P.; Barieau, R. E.; Donohue; Lu, J. C.-S. The crystal structure of DL-serine. *Acta Crystallogr.* **1953**, *6*, 241−256.
(83) Kistenmacher, T. J.; Rand, G. A.; Marsh, R. E. Refinements of the crystal structures of DL-serine and anhydrous L-serine. *Acta Crystallogr. B* **1974**, *30*, 2573−2578.
(84) Frey, M. N.; Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. XI. Molecular configuration and hydrogen bonding of serine in the crystalline amino acids L-serine monohydrate and DL-serine. *Acta Crystallogr. B* **1973**, *29*, 876−884.
(85) Dittrich, B.; Hubschle, C. B.; Messerschmidt, M.; Kalinowski, R.; Girnt, D.; Luger, P. The invariom model and its application: refinement of D,L-serine at different temperatures and resolution. *Acta Crystallogr.A* **2005**, *61*, 314−320.
(86) Boldyreva, E. V.; Kolesnik, E. N.; Drebushchak, T. N.; Sowa, H.; Ahsbahs, H.; Seryotkin, Y. V. A comparative study of the anisotropy of lattice strain induced in the crystals of DL-serine by cooling down to 100 K, or by increasing pressure up to 8.6 GPa. A comparison with L-serine. *Z. Kristallogr.* **2006**, *221*, 150−161.
(87) Benedetti, E.; Pedone, C.; Sirigu, A. The crystal structure of L-(−)-serine. *Gazz. Chim. Ital.* **1973**, *103*, 555−561.
(88) Moggach, S. A.; Allan, D. R.; Morrison, C. A.; Parsons, S.; Sawyer, L. Effect of pressure on the crystal structure of L-serine-I and the crystal structure of L-serine-II at 5.4 GPa. *Acta Crystallogr. B.* **2005**, *61*, 58.
(89) Boldyreva, E. V.; Kolesnik, E. N.; Drebushchak, T. N.; Ahsbahs, H.; Beukes, J. A.; Weber, H.-P. A comparative study of the anisotropy of lattice strain induced in the crystals of L-serine by cooling down to 100 K or by increasing pressure up to 4.4 GPa. *Z. Kristallogr.* **2005**, *220*, 58−65.
(90) Drebushchak, T. N.; Sowa, H.; Seryotkin, Y. V.; Boldyreva, E. V.; Ahsbahs, H. L-Serine III at 8.0 GPa. *Acta Crystallogr. E* **2006**, *62*, o4052−o4054.
(91) Boldyreva, E. V.; Sowa, H.; Seryotkin, Y. V.; Drebushchak, T. N.; Ahsbahs, H.; Chernyshev, V.; Dmitriev, V. Pressure-induced phase transitions in crystalline l-serine studied by single-crystal and high-resolution powder X-ray diffraction. *Chem. Phys. Lett.* **2006**, *429*, 474−478.
(92) Ramanadham, M.; Sikka, S. K.; Chidambaram, R. Structure determination of Ls-theonine by neutron diffraction. *Pramana* **1973**, *1*, 247−259.
(93) Janczak, J.; Zobel, D.; Luger, P. L-Threonine at 12 K. *Acta Crystallogr. C* **1997**, *53*, 1901−1904.
(94) Sequeira, A.; Rajagopal, H.; Ramanadham, H. A neutron study on the structure of dl-aspartic acid. *Acta Crystallogr. C* **1989**, *45*, 906−908.
(95) Flaig, R.; Koritsánszky, T.; Zobel, D.; Luger, P. Topological Analysis of the Experimental Electron Densities of Amino Acids. 1. d,l-Aspartic Acid at 20 K. *J. Am. Chem. Soc.* **1998**, *120*, 2227−2238.
(96) Rao, S. T. Refinement of DL-aspartic acid. *Acta Crystallogr. B* **1973**, *29*, 1718−1720.
(97) Wang, G.-M.; Li, Z.-X.; Duan, C.-S.; Li, H. DL-Aspartic acid. *Acta Crystallogr. E* **2007**, *63*, o4003.
(98) Derissen, J. L.; Endeman, H. J.; Peerdeman, A. F. The crystal and molecular structure of l-aspartic acid. *Acta Crystallogr. B,.* **1968**, *24*, 1349−1354.
(99) Zhang, Y.-M.; Gao, C.; Lin, Q.; Yang, L.-Z.; Wei, T.-B. *Huaxue Yanjiu Yu Yingyong (Chin.) (Chem. Res. Appl.)* **2006**, *18*, 952
(100) Bendeif, E.; Jelsch, C. The experimental library multipolar atom model refinement of L-aspartic acid. *Acta Crystallogr. C* **2007**, *63*, o361−o364.
(101) Marcoin, W.; Duda, H.; Kusz, J.; Bzowski, B.; Warczewski, J. Structure rerefinement of L-glutamic acid with the aid of the X-ray 4-circle diffractometer, *Proceedings of the 17th Applied Crystallography Conference*, 1999; p 40, World Science: Wisła, Poland 1998/03.
(102) Hirayama, N.; Shirahata, K.; Ohashi, Y.; Sasada, Y. Structure of α Form of L-Glutamic Acid. α -β Transition. *Bull. Chem. Soc. Jpn.* **1980**, *53*, 30−35.
(103) Lehmann, M. S.; Nunes, A. C. A short hydrogen bond between near identical carboxyl groups in the α -modification of L-glutamic acid. *Acta Crystallogr. B* **1980**, *36*, 1621−1625.

(104) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. VIII: the crystal and molecular structure of the β-form of the amino acidl -glutamic acid. *J. Cryst. Mol. Struct.* **1972**, *2*, 225−233.
(105) Koetzle, T. F.; Frey, M. N.; Lehmann, M. S.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. XIII. Molecular and crystal structure of the amino acid L-glutamine. *Acta Crystallogr. B* **1973**, *29*, 2571−2575.
(106) Wagner, A.; Luger, P. Charge Density and Topological Analysis of L-Glutamine. *J. Mol. Struct.* **2001**, *595*, 39−46.
(107) Suresh, S.; Padmanabhan, S.; Vijayan, M. DL-Glutamine. *Acta Crystallogr. C* **1996**, *52*, 1313−1316.
(108) Yamada, K.; Hashizume, D.; Shimizu, T.; Yokoyama, S. L-Asparagine. *Acta Crystallogr. E* **2007**, *63*, o3802−o3803.
(109) Hübschle, C. B.; Messerschmidt, M.; Luger, P. Crystal structure of DL-Tryptophan at 173K. *Cryst. Res. Technol.* **2004**, *39*, 274−278.
(110) Edington, P.; Harding, M. M. The crystal structure of DL-histidine. *Acta Crystallogr. B* **1974**, *30*, 204−206.
(111) Coppens, P.; Abramov, Y.; Carducci, M.; Korjov, B.; Novozhilova, I.; Alhambra, C.; Pressprich, M. R. Experimental charge densities and intermolecular interactions: electrostatic and topological analysis of DL-histidine. *J. Am. Chem. Soc.* **1999**, *121*, 2585−2593.
(112) Madden, J. J.; McGandy, E. L.; Seeman, N. C.; Harding, M. M.; Hoy, A. The crystal structure of the monoclinic form of l-histidine. *Acta Crystallogr. B* **1972**, *28*, 2382−2389.
(113) Averbuch-Pouchot, M. T. Crystal structure of l-histidinium phosphite and a structure reinvestigation of the monoclinic form of l-histidine. *Z. Kristallogr.* **1993**, *207*, 111−120.
(114) Madden, J. J.; McGandy, E. L.; Seeman, N. C. The crystal structure of the orthorhombic form of L-(+)-histidine. *Acta Crystallogr. B* **1972**, *28*, 2377−2382.
(115) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. Precision Neutron Diffraction Structure Determination of Protein and Nucleic Acid Components. IV. The Crystal and Molecular Structure of the Amino Acid L-Histidine. *Int. J. Pept. Protein Res.* **1972**, *4*, 229−239.
(116) Dunitz, J. D.; Schweizer, W. B. Anhydrous DL-Glutamic Acid. *Acta Crystallogr. C* **1995**, *51*, 1377−1379.
(117) Dobrowolski, J. Cz.; Rode, J. E.; Sadlej, J. Cysteine conformations revisited. *J. Mol. Struct. (Theochem)* **2007**, *810*, 129−134.
(118) Dobrowolski, J. Cz.; Jamróz, M. H.; Kołos, R.; Rode, J. E.; Sadlej, J. Theoretical prediction and the first IR-matrix observation of several L-cysteine molecule conformers. *ChemPhysChem.* **2007**, *8*, 1085−1094.
(119) Sadlej, J.; Dobrowolski, J. Cz.; Rode, J. E.; Jamróz, M. H. Density Functional Theory Study on Vibrational Circular Dichroism as a Tool for Analysis of Intermolecular Systems: (1:1) Cysteine-Water Complex Conformations. *J. Phys. Chem. A* **2007**, *111*, 10703−10711.
(120) Dobrowolski, J. Cz.; Jamróz, M. H.; Kołos, R.; Rode, J. E.; Sadlej, J. IR low-temperature matrix and *ab initio* study on β-alanine conformers. *ChemPhysChem* **2008**, *9*, 2042−2051.
(121) Dobrowolski, J. Cz.; Jamróz, M. H.; Kołos, R.; Rode, J. E.; Cyrański, M. K.; Sadlej, J. IR low-temperature matrix, X-ray and ab initio study on L-isoserine conformations. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10818−10830.
(122) Rode, J. E.; Dobrowolski, J. Cz.; Sadlej, J. Phenylisoserine in the gas-phase and water: Ab initio studies on neutral and zwitterion conformers. *J. Mol. Model.* **2011**, *17*, 961−970.
(123) Boon, G.; van Alsenoy, C.; de Proft, F.; Bultinck, P.; Geerlings, P. Similarity and Chirality: Quantum Chemical Study of Dissimilarity of Enantiomers. *J. Phys. Chem. A* **2003**, *107*, 11120−11127.
(124) Boon, G.; van Alsenoy, C.; de Proft, F.; Bultinck, P.; Geerlings, P. Molecular quantum similarity of enantiomers of amino acids: a case study. *J. Mol. Struct. Theochem* **2005**, *727*, 49−56.
(125) Boon, G.; van Alsenoy, C.; de Proft, F.; Bultinck, P.; Geerlings, P. Study of Molecular Quantum Similarity of Enantiomers of Amino Acids. *J. Phys. Chem. A* **2006**, *110*, 5114−5120.
(126) Geerlings, P.; Boon, G.; van Alsenoy, C.; de Proft, F. Density Functional Theory and Quantum Similarity. *Int. J. Quantum Chem.* **2005**, *101*, 722−732.
(127) Carbó, R.; Arnau, M.; Leyda, L. How similar is a molecule to another? *Int. J. Quantum Chem.* **1980**, *17*, 1185.
(128) Bultinck, P.; Gironés, X.; Carbó-Dorca, R. Molecular Quantum Similarity: Theory and Applications. *Rev. Comput. Chem.* **2005**, *21*, 127−207.
(129) Gironés, X.; Robert, D.; Carbó-Dorca, R. TGSA:a molecular superposition program based on Topo-Geometrical Considerations. *J. Comput. Chem.* **2001**, *22*, 255−263.
(130) Dryzun, C.; Avnir, D. Chirality Measures for Vectors, Matrices, Operators and Functions. *ChemPhysChem* **2011**, *12*, 197−205.
(131) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.
(132) Tropsha, A. Recent Advances in Development, Validation, and Exploitation of QSAR Models. In *Burger's Medicinal Chemistry and Drug Discovery*, 7th ed.; Abraham, D., Ed.; John Wiley & Sons, Inc.: New York, 2010; Vol. *1*, pp 505−533.
(133) Cherkasov, A.; Ban, F.; Santos-Filho, O.; Thorsteinson, N.; Fallahi, M.; Hammond, G. L. An Updated Steroid Benchmark Set and Its Application in the Discovery of Novel Nanomolar Ligands of Sex Hormone-Binding Globulin. *J. Med. Chem.* **2008**, *51*, 2047−2056.
(134) Fang, H.; Tong, W.; Branham, W. S.; Moland, C. L.; Dial, S. L.; Hong, H.; Xie, Q.; Perkins, R.; Owens, W.; Sheehan, D. M. Study of 202 Natural, Synthetic, and Environmental Chemicals for Binding to the Androgen Receptor. *Chem. Res. Toxicol.* **2003**, *16*, 1338−1358.
(135) de Jésus-Tran, K. P.; Côté, P. L.; Cantin, L.; Blanchet, J.; Labrie, F.; Breton, R. Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. *Protein Sci.* **2006**, *15*, 987−999.
(136) Keinan, S.; Avnir, D. Quantitative Chirality in Structure-Activity Correlations. Shape Recognition by Trypsin, by the D2 Dopamine Receptor, and by Cholinesterases. *J. Am. Chem. Soc.* **1998**, *120*, 6152−6159.
(137) Keinan, S.; Avnir, D. Quantitative Symmetry in Structure-Activity Correlations: The Near C2 Symmetry of Inhibitor/HIV Protease Complexes. *J. Am. Chem. Soc.* **2000**, *122*, 4378−4384.
(138) Katzenelson, O.; Edelstein, J.; Avnir, D. Quantitative chirality of helicenes. *Tetrah. Asym.* **2000**, *11*, 2695−2704.
(139) Pogliani, L. Molecular Connectivity Model for Determination of Physicochemical Properties of α-Amino Acids. *J. Phys. Chem.* **1993**, *97*, 6731−6736.
(140) Randić, M; Dobrowolski, J. Cz. Optimal Molecular Connectivity Descriptors for Nitrogen-Containing Molecules. *Int. J. Quantum Chem.* **1998**, *70*, 1209−1215.
(141) Randić, M; Milles, D.; Basak, S. C. On Characterization of Physical Properties of Amino Acids. *Int. J. Quantum Chem.* **2000**, *80*, 1199−1209.
(142) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(143) Estrada, E. Physicochemical Interpretation of Molecular Connectivity Indices. *J. Phys. Chem. A* **2002**, *106*, 9085−9091.
(144) Pogliani, L. Structure property relationships of amino acids and some dipeptides. *Amino Acids* **1994**, *6*, 141−153.
(145) Pogliani, L. Molecular Connectivity Descriptors of the Physicochemical Properties of the α-Amino Acids. *J. Phys. Chem.* **1994**, *98*, 1494−1499.
(146) Pogliani, L. Modeling the solubility and activity of amino acids with the LCCI method. *Amino Acids* **1995**, *9*, 217−228.
(147) Sirimulla, S.; Lerma, M.; Herndon, W. C. Prediction of Partial Molar Volumes of Amino Acids and Small Peptides: Counting Atoms versus Topological Indices. *J. Chem. Inf. Model.* **2010**, *50*, 194−204.
(148) Miličević, A.; Raos, N. Estimation of Stability Constants of Copper(II) Complexes with α-Amino Acids Using Connectivity Index $^3\chi^v$. Common Model for the Binary and Ternary Complexes. *Chin. J. Chem.* **2011**, *29*, 1800−1804.

(149) Miličević, A.; Raos, N. A model to estimate stability constants of amino acid chelates with Cu(II) and Ni(II) at different ionic strengths. *J. Mol. Liq.* **2012**, *165*, 139−142.
(150) Pogliani, L. Modeling of Properties of Amino Acids with Random, Semirandom, and Molecular Connectivity Descriptors. *Int. J. Quantum Chem.* **2012**, *112*, 2267−2274.