

Competition of hydrogen-bond acceptors for the strong  
carboxyl donor

Thomas Steiner

Institut für Chemie – Kristallographie,  
Freie Universität Berlin, Takustrasse 6,  
D-14195 Berlin, GermanyCorrespondence e-mail:  
steiner@chemie.fu-berlin.de

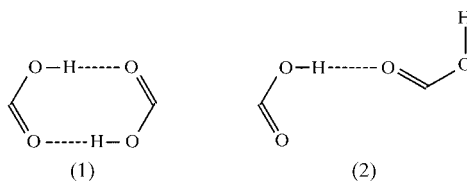
A database study on the competition of hydrogen-bond acceptors for the strong carboxyl donor in crystals is reported. The relative success in attracting this donor is determined for 34 types of acceptor (O, N, S, halogen and  $\pi$  acceptors), and a correlation between the success in competition and the average hydrogen-bond distance is established.

Received 12 September 2000

Accepted 13 October 2000

## 1. Introduction

Carboxyl groups are among the best investigated hydrogen-bond functionalities (Jeffrey, 1997). Since they possess a hydrogen-bond donor as well as an acceptor site, carboxyl groups can readily form hydrogen bonds between each other as cyclic dimers (1) or open arrays such as (2). The roles of hydrogen-bonding carboxyl groups in crystal packing have been analyzed in depth (reviewed by Bernstein *et al.*, 1994) and they have become a popular building block in deliberate crystal design ('crystal engineering'; *e.g.* Desiraju, 1989; MacDonald & Whitesides, 1994).



By intuition, one is tempted to take the carboxyl dimer as a very stable and robust hydrogen-bond pattern. The practising crystallographer, on the other hand, knows by experience that the formation of this motif is not very reliable and is, in competitive situations, often displaced by other hydrogen-bond configurations. In an important statistical analysis, Allen *et al.* (1999) determined the global probabilities of formation of 75 bimolecular hydrogen-bonded ring motifs in organic crystal structures. For the carboxyl dimer, they found a probability of formation of only 33% averaged over all competitive situations and this relatively low probability is explained by competition with other hydrogen-bond patterns. For a better understanding of the mechanisms governing the make-up of hydrogen-bond arrays, quantification of these competitive effects is necessary. In the present communication, a brief database study is reported on the competition of different

acceptor types for the strong carboxyl donor. The aim is to quantify the success rates of a wide variety of acceptors, ranging from very strong to very weak.

## 2. Data retrieval

Update 5.19 of the CSD (Cambridge Structural Database; Allen & Kennard, 1993) was used (April 2000, 215 403 entries). A subset was created containing 2765 good quality crystal structures of carboxylic acids, which served as the basis for all further searches (only ordered and error-free crystal structures with  $R < 0.07$ , organic as well as organometallic, acid H-atom located; a number of structures with obviously unrealistic H-atom positions had to be excluded). H-atom positions were normalized using default parameters.

Since carboxyl groups are strong proton donors, relatively stringent geometric cutoff definitions could be selected for identifying hydrogen bonds ( $H \cdots A < 2.2 \text{ \AA}$  for  $A = O, N, F$ ;  $< 2.8 \text{ \AA}$  for  $A = S, Cl$ ;  $< 3.0 \text{ \AA}$  for  $A = Br$ ;  $< 3.2 \text{ \AA}$  for  $A = \pi$ ;  $O-H \cdots A$  in all cases  $> 130^\circ$ ). Of bifurcated hydrogen bonds, only the shorter component was considered. Standard uncertainties of the relative frequencies given in Table 1,  $h(x)$ , were estimated as  $s.u.(h) = [h(1-h)/n]^{1/2}$ . Structure illustrations were drawn using *PLUTON* (Spek, 1995).

## 3. Results and discussion

In the data set containing 3617 carboxyl groups, a total of 1039 hydrogen bonds between carboxyl functions are identified. Of these, 881 (= 85%) constitute the carboxyl dimer (1) and the remaining 158 (= 15%) form non-cyclic arrangements, see (2). The other 2578 carboxyl functions donate hydrogen

**Table 1**

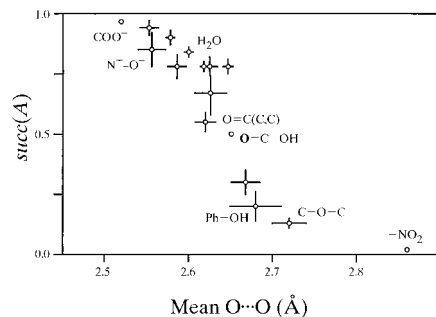
Hydrogen bonds  $\text{O}=\text{C}-\text{OH}\cdots A$  in carboxylic acid structures where acceptors  $A$  are present, irrespective of the presence of further potential acceptors  $A'$ .

$n[(1) + (2)]$  = number of hydrogen bonds between carboxyl groups, sum of motifs (1) and (2).  $\langle \text{O}\cdots A \rangle$  is the mean  $\text{O}\cdots A$  distance in Å. 'Relative success' of  $A$ :  $\text{succ}(A) = n(\text{OH}\cdots A)/[n(\text{OH}\cdots A) + n[(1) + (2)]]$ .  $M$  = transition metal atom.

Acceptor $A$	$n(\text{struct.}/A/\text{COOH})$	$n(\text{OH}\cdots A)$	$n[(1) + (2)]$	$\langle \text{O}\cdots A \rangle$	Success of $A$
<b>Oxygen acceptors</b>					
$\text{COO}^-$	514/622/658	561	20	2.520 (3)	0.966 (8)
$\text{O}=\text{C}-\text{O}-M$	95/180/136	61	4	2.553 (11)	0.94 (3)
$\text{P}=\text{O}, \text{P}-\text{O}^-$	89/116/103	86	10	2.578 (5)	0.90 (3)
$\text{N}^+-\text{O}^-$	23/28/29	23	4	2.556 (16)	0.85 (7)
$\text{H}_2\text{O}$	466/802/720	277	53	2.600 (5)	0.84 (2)
$\text{O}=\text{C}(\text{N},\text{N})$	84/98/107	50	14	2.586 (11)	0.78 (5)
$\text{O}=\text{C}(\text{C},\text{N})$	440/641/525	231	64	2.618 (4)	0.78 (2)
$\text{S}=\text{O}, \text{S}-\text{O}^-$	148/414/199	98	27	2.625 (9)	0.78 (4)
$\text{C}(\text{sp}^3)-\text{OH}$	408/780/514	190	52	2.647 (6)	0.78 (3)
$M-\text{O}-X$	105/207/150	18	9	2.626 (19)	0.67 (9)
$\text{O}=\text{C}(\text{C},\text{C})$	221/258/260	84	69	2.620 (12)	0.55 (4)
$\text{O}=\text{C}-\text{OH}$	2765/3617/3617	1039	1039	2.651 (1)	0.500
$\text{O}=\text{C}-\text{O}-\text{C}$	173/265/199	21	50	2.668 (17)	0.30 (5)
$\text{Ph}-\text{OH}$	110/154/130	9	37	2.68 (3)	0.20 (6)
$\text{C}-\text{O}-\text{C}$	595/1021/736	35	226	2.72 (2)	0.13 (2)
$-\text{NO}_2$	132/259/167	1	65	2.86	0.02 (2)
$\text{N}/\text{P}/\text{S}-\text{OH}$	81/108/95	0	15	-	0.0
$M-\text{CO}$	26/97/31	0	22	-	0.0
<b>Nitrogen</b>					
Pyridyl $\text{N}\ddagger$	185/288/266	156	16	2.669 (5)	0.91 (2)
$\text{C}=\text{N}-\text{O}$	21/29/29	16	4	2.706 (9)	0.80 (9)
$\text{C}=\text{N}-\text{N}\ddagger$	47/72/62	12	11	2.695 (12)	0.52 (10)
$-\text{C}\equiv\text{N}$	29/52/37	5	14	2.74 (4)	0.26 (10)
$\text{C}(\text{sp}^2)-\text{NH}_2$	170/266/217	0	31	-	0.0
<b>Sulfur</b>					
$\text{C}=\text{S}$	28/37/32	4	11	3.17 (5)	0.27 (11)
$\text{C}-\text{S}-\text{C}$	119/148/136	0	47	-	0.0
<b>Halogen</b>					
$\text{F}^-$	4/4/5	5	0	2.468 (15)	1.0
$\text{Cl}^-$	125/154/156	101	3	2.998 (5)	0.971 (16)
$\text{Br}^-$	26/30/30	17	4	3.16 (2)	0.81 (9)
$\text{Cl}-M$	54/103/75	5	20	3.10 (3)	0.20 (8)
$\text{F}-\text{C}$	105/260/126	0	46	-	0.0
$\text{Cl}-\text{C}$	198/394/236	0	121	-	0.0
$\text{Br}-\text{C}$	29/41/34	0	15	-	0.0
<b><math>\pi</math> acceptors</b>					
$\text{Ph}$	1391/2573/1792	0	703	-	0.0
$\text{C}=\text{C}$	672/1117/846	0	317	-	0.0

‡ Only uncharged systems considered.

bonds to a great variety of acceptors. In the following, arrays (1) and (2) are no longer distinguished, but merged to a single group


**Figure 1**

Hydrogen bonds from carboxyl groups to O-atom acceptors. Correlation of the mean  $\text{O}\cdots\text{O}$  distance and the relative success in competing with hydrogen bonding between carboxyl groups,  $\text{succ}(A) = n(\text{OH}\cdots A)/[n(\text{OH}\cdots A) + n[(1) + (2)]]$ . Numerical values are given in Table 1.

(1) + (2), representing any kind of hydrogen bond between carboxyl groups.

The competition of (1) + (2) with other hydrogen-bond arrays can be considered as a competition between the carboxylic acceptor and other acceptors  $A$  for the strong acid donor. The situation can be analyzed in different ways. A stringent method would be to retrieve crystal structures that contain carboxyl groups and a particular competitor  $A$  in equal numbers, but no other competitors  $A'$ . The carboxyl functions in these structures would donate only two types of hydrogen bond, (1) + (2) and  $\text{O}-\text{H}\cdots A$ , and the relative frequencies would adequately represent the success in competition. Unfortunately, this method yields only small or even null data samples for most acceptors  $A$  and cannot be used for a general analysis. Therefore, a rougher method had to be used here. Carboxylic acid

structures were retrieved that contain a particular acceptor  $A$ , irrespective of the absolute numbers of carboxyl functions and  $A$ , and irrespective of the presence of further acceptor types  $A'$ . The carboxyl functions in such a sample donate hydrogen bonds (1) + (2),  $\text{O}-\text{H}\cdots A$  and  $\text{O}-\text{H}\cdots A'$ . A 'relative success'  $\text{succ}(A)$  of  $\text{O}-\text{H}\cdots A$  hydrogen bonds in the competition with (1) + (2) can be quantified as  $\text{succ}(A) = n(\text{OH}\cdots A)/[n(\text{OH}\cdots A) + n[(1) + (2)]]$ . For hydrogen-bond types that are favoured over (1) + (2),  $\text{succ}(A)$  is in the range 0.5–1.0 and for those that are disfavoured it is 0.0–0.5. A relative success of  $\text{succ}(A) = 0.8$ , for example, means that the  $\text{O}-\text{H}\cdots A$  hydrogen bonds in the sample outnumber (1) + (2) by a factor of four. The competition of (1) + (2) as well as of  $\text{O}-\text{H}\cdots A$  with the other acceptors,  $\text{O}-\text{H}\cdots A'$ , does not show up in  $\text{succ}(A)$ , but only in the absolute numbers  $n[(1) + (2)]$ ,  $n(\text{OH}\cdots A)$  and  $n(\text{OH}\cdots A')$ .

Using this method, success in attracting the carboxyl donor relative to the formation of (1) + (2) was determined for 34 types of O, N, S, Hal and  $\pi$  acceptors, Table 1. Carboxyl groups are also found to donate hydrogen bonds to other groups (such as  $\text{NO}_3^-$ ,  $M=\text{O}$ ,  $\text{As}=\text{O}$ ,  $\text{N}=\text{N}=\text{N}$ ,  $\text{I}^-$  etc.), but for these the samples are too small to allow statistically meaningful analysis. Surprisingly, most types of O and N acceptors are more successful in competing for the strong donor than the carboxyl group itself. Most successful are carboxylate anions with  $\text{succ}(\text{COO}^-) = 0.966$  (8), but the weaker water and hydroxyl acceptors also perform quite well, with  $\text{succ}(\text{H}_2\text{O}) = 0.84$  (2) and  $\text{succ}(\text{Csp}^3-\text{OH}) = 0.78$  (4). Phenol and ether O acceptors have values of  $\text{succ}(A)$  clearly below 0.5, but are still strong enough to attract carboxylic acid donors with appreciable frequencies. The weak nitro acceptor is found only once in the CSD involved in a hydrogen bond with a carboxyl group, leading to a value of  $\text{succ}(\text{NO}_2) = 0.02$  (CSD code ZAJHEX; Smith *et al.*, 1995), and carbonyl ligands of organometallic compounds have as yet never been observed in that role (even though they can accept hydrogen bonds in principle).

Of the non-O acceptors, pyridyl N atoms and halide ions are particularly successful in competing for carboxyl donors [ $\text{succ}(\text{N}_{\text{py}}) = 0.91$  (2),  $\text{succ}(\text{Cl}^-) = 0.97$  (2)], but also the relatively weak acceptors  $-\text{C}\equiv\text{N}$  and  $\text{C}=\text{S}$  succeed better than might be assumed, with  $\text{succ}(A)$  values around 0.25. No hydrogen bonds with carboxyl donors are found in the CSD for the weak acceptors  $\text{Hal}-\text{C}$ ,  $\pi(\text{Ph})$  and  $\pi(\text{C}=\text{C})$ ; with weaker  $\text{O}-\text{H}$  donors

such as H<sub>2</sub>O, these groups do occasionally accept hydrogen bonds (Desiraju & Steiner, 1999).

Table 1 lists the mean O···A distances, (O···A), for the different hydrogen bond types, and it is of interest to see if there is a correlation between this distance and *succ*(A). In Fig. 1, *succ*(A) is drawn for O acceptors against (O···O), and shows a clear S-shaped correlation. The most successful acceptor, COO<sup>-</sup>, also has the shortest average distance in hydrogen bonds, 2.520 (3) Å, and with increasing average distance the success in competition with (1) + (2) gradually falls. For N acceptors, Table 1 contains too few acceptor types to construct a diagram such as that given in Fig. 1, but nevertheless it is obvious that the trend is the same as for O: shorter hydrogen bonds have better chances of being formed in competitive situations. This supports the

classical length–strength relation of hydrogen bonds; to avoid misunderstanding, however, it must be stressed that this relation is only valid for *mean values* of large structure samples, and because of large scatter within the samples, it is not generally applicable for individual cases (discussed e.g. by Desiraju & Steiner, 1999).

Very important features of Fig. 1 are (a) the smooth and continuous nature of the correlation and (b) the fact that a value of *succ*(A) = 1.0 is not reached even for the strongest acceptors. The latter point means that even very strong acceptors such as carboxylates and phosphates do not attract a carboxyl donor ‘with certainty’, but there is a residual chance of formation for (1) + (2). Analogously, on the other end of the correlation, even quite weak acceptors have a non-zero chance of binding a carboxyl donor. It may also be assumed that for some

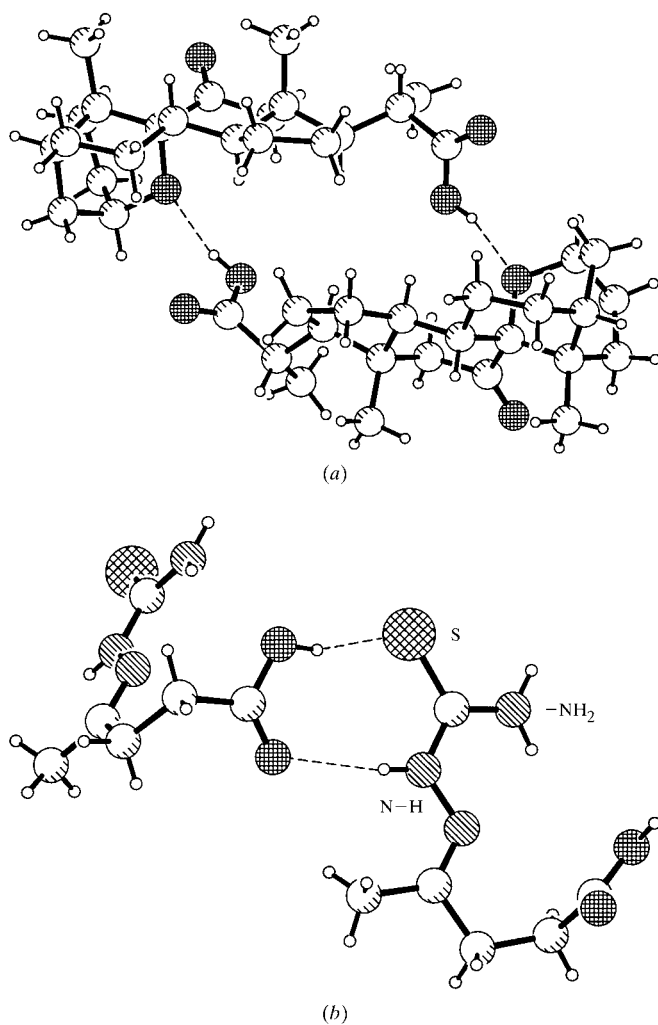
of the weak acceptors with *succ*(A) = 0 in Table 1, larger structure samples might contain a few hydrogen bonds.

Since it is an essential point that weak acceptors do occasionally bind strong donors, this is discussed further, and illustrated by two examples in Fig. 2. When browsing through such structures displayed by a suitable CSD query, it becomes obvious that most examples belong to a small number of overall situations. Most frequent is the situation where the carboxyl group is carried by an awkwardly shaped molecule. Then, packing considerations may prevent the formation of favourable hydrogen-bond motifs and ‘second choice’ types of hydrogen bonds are formed as substitutes. A typical example with a steroid molecule is shown in Fig. 2(a), where the carboxyl groups donate hydrogen bonds to ether-type acceptors (note also that carbonyl acceptors are present). A different situation arises if the carboxyl group is involved in a hydrogen-bond array (‘supramolecular synthon’; Desiraju, 1995) that is for some reason favourable, despite the participation of a weak acceptor. An example for this case is shown in Fig. 2(b), where a cyclic motif is formed by an O–H···S=C and an N–H···O=C hydrogen bond (note the structural relation with (1); see Bernstein *et al.*, 1995). Finally, carboxyl groups sometimes donate hydrogen bonds to weak acceptors if these simply are present in great numerical surplus; this is a relatively trivial but not infrequent case.

The observations presented here show (once more) that the competitive situation involving hydrogen-bond donors and acceptors of different strengths is subtle and complicated. Even though it is true that strong hydrogen-bond donors tend to interact with strong acceptors, this is valid only as a *tendency*. Weak acceptors also have a certain chance of attracting the strong donor. This weakens the general applicability of rules for predicting hydrogen-bond modes from hierarchies of donor and acceptor strengths and indeed all such rules published are very unreliable in practice.

## APPENDIX A Note concerning disorder

A referee of this paper raised an interesting question concerning the methodology of data retrieval. The carboxyl dimer (1) is known to have a tendency for orientational disorder with partially occupied H-atom positions. Disordered structures have been excluded from the analysis and one could



**Figure 2**  
Examples of carboxyl donors hydrogen bonding with relatively poor acceptors. (a) With a C–O–C acceptor in (+)-23,24-dinor-3 $\alpha$ ,9 $\alpha$ -epoxy-11-oxo-5 $\beta$ -cholan-22-oic acid, O···O = 2.70 and 2.72 Å (structure published by Thompson *et al.*, 1999; CSD CATQOD). (b) With an S=C acceptor in 2-[1-(2-carboxyethyl)ethylidene]-hydrazinecarbothioamide, O···S = 3.09 Å (structure published by Ng, 1992; CSD JUBMAU).

assume that this criterion for exclusion disfavors structures containing (1) to a larger degree than the others. This would lead to bias in the values of  $succ(A)$ , which would be determined too high. The referee suggests determining  $succ(A)$  values for a data sample also containing disordered structures, using a O...O distance criterion for identification of hydrogen bonds. To clear this point, such an analysis was actually performed for a test set of three acceptor types (structures used with  $R < 0.07$ , irrespective of disorder and if H atoms have been located at all; stringent O...O distance cutoff at 3.0 Å). For two of the three test cases, the new values  $occ(A)$  are actually slightly smaller than those in Table 1, but the difference is around or smaller than one standard uncertainty [ $succ(\text{H}_2\text{O}) = 0.82$  (2) versus 0.84 (2) in Table 1;  $succ(\text{O}=\text{C}_{\text{C,N}}) =$

0.78 (2) versus 0.78 (2) in Table 1;  $succ(\text{Ph}-\text{OH}) = 0.18$  (5) versus 0.20 (6) in Table 1]. This means that an effect of disorder phenomena on the values in Table 1 is only very minor.

The author thanks Professor Wolfram Saenger for the opportunity to carry out this study in his laboratory.

### References

- Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 1.
- Allen, F. H., Motherwell, W. D. S., Raithby, P. R., Shields, G. P. & Taylor, R. (1999). *New J. Chem.* **23**, 25–34.
- Bernstein, J., Davis, R. E., Shimoni, L. & Chang, N.-L. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555–1573.
- Bernstein, J., Etter, M. C. & Leiserowitz, L. (1994). *Structure Correlation*, edited by H.-B. Bürgi & J. D. Dunitz, Vol. 2, pp. 431–507. Weinheim: VCH.
- Desiraju, G. R. (1989). *Crystal Engineering. The Design of Organic Solids*. Amsterdam: Elsevier.
- Desiraju, G. R. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 2311–2327.
- Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*. Oxford University Press, Oxford.
- Jeffrey, G. R. (1997). *An Introduction to Hydrogen Bonding*. Oxford University Press.
- MacDonald, J. C. & Whitesides, G. M. (1994). *Chem. Rev.* **94**, 2383–2420.
- Ng, S. W. (1992). *Acta Cryst.* **C48**, 2057–2058.
- Smith, G., Lynch, D. E., Byriel, K. A. & Kennard, C. H. (1995). *Aust. J. Chem.* **48**, 1133–1149.
- Spek, A. L. (1995). *PLUTON*. University of Utrecht, The Netherlands.
- Thompson, H. W., Brunskill, A. P. J. & Lalancete, R. A. (1999). *Acta Cryst.* **C55**, 1677–1680.