

Introduction

Hydrogen bonds are generally recognized as vital to biopolymer structures. Yet, there are still various controversial issues. One of them is the importance of C-H...X interactions (X: O,N). Due to the **rapidly increasing number of new structures** we are now able to proceed from discussions of single interactions and single structures to the identification of **general patterns and trends in large structure sets**. For a small dataset of 13 structures Derewenda et al. (*J. Mol. Biol.* 1995, 252, 248) have analyzed C-H...O interactions in proteins. We have scanned large sets of RNA, DNA and protein structures for potential hydrogen bonds. Here we report on **C-H...X interactions in RNA structures**. Intra-residue C-H...O contacts in nucleic acids between the purine C8-H or pyrimidine C6-H and O5' are formed when the bases assume anti-conformation (Rubin et al., *Biochemistry* 1972, 11, 312). A short C-H...O contact was described between C5 and O4 in an U-U base pair (Calcutta pair; Wahl et al., *Nature Struct. Biol.* 1996, 3, 24). In molecular dynamics simulations on the anticodon loop of tRNA(ASP) stable trajectories were found for the nucleotide pairs U33-C36 and U33-U35 (Auffinger et al., *J. Am. Chem. Soc.* 1996, 118, 1181). Apart from these single types of interactions no information on C-H...X interactions was available up to now. Our aim was to identify all short C-H...X contacts in RNA and to answer the question whether or not they can be regarded as hydrogen bonds.

Methods

The following RNA structure sets have been selected from the Protein Data Bank (PDB):

- I high-resolution X-ray structures, resolution < 2 Å; 5 structures:**
157d, 1osu, 1rxh, 259d, 1urn;
- II X-ray, 23 structures:**
157d, 1gid, 1mme, 1osu, 1rna, 1rxa, 1rxh, 1sdr, 1tn2, 1tra, 205d, 255d, 259d, 280d, 283d, 299d, 2tra, 300d, 301d, 3tra, 4tna, 4tra, 4tna;
- III X-ray RNA-protein complexes; 10 structures:**
1asy, 1asz, 1gr, 1igs, 1hrs, 1qr, 1qr, 1qr, 1qr, 1qr, 1qr;
- IV NMR; 13 structures, 69 models:**
1arj (20), 1elh (6), 1fmm (5), 1hbx (20), 1kaj, 1koc, 1kod, 1rau, 1rht, 1rng (5), 1rnk, 1scl (6), 1slo;
- V all structures.**

Only inter-residue interactions were considered. Interactions have been identified by the program HBexplore (Lindauer, Bendic, Sühnel, *Comput. Appl. Biosci.* 1996, 12, 281; online version of the paper: http://www.imb-jena.de/www_bioc/pub/pubhbexplore.pdf; program homepage: http://www.imb-jena.de/www_bioc/hbx/hbx.html).

Hydrogen atoms have been added to X-ray structures according to standard geometrical rules. Those calculated hydrogen atoms have also been used in the analysis of the NMR structures.

All C-H...X contacts with hydrogen-acceptor distances of smaller than 3.5 Å have been taken into account. For part of the analysis the following angle cutoffs have been used:

angle CHX > 90°, angle CXX1 > 90°, angle HXX1 > 90°

(C: H-bond donor; X: H-bond acceptor; X1: covalent neighbour of X). The number of hydrogen bonds identified has been normalized with regard to the size of volume elements and is given as number of H-bonds *1000/Å³.

Results and Discussion

C-H...X contacts in different structure sets

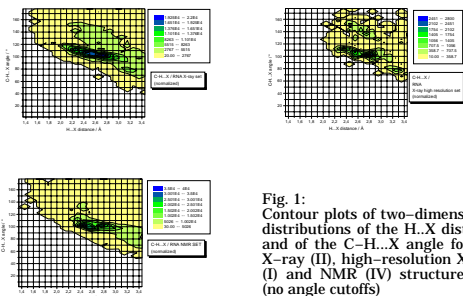


Fig. 1: Contour plots of two-dimensional distributions of the H...X distance and of the C-H...X angle for the X-ray (II), high-resolution X-ray (I) and NMR (IV) structure sets (no angle cutoffs)

The geometrical distributions of C-H...X contacts are similar for NMR and X-ray structures (Fig. 1). Therefore, we have used the total set V, which includes both structures determined by X-ray diffraction and NMR spectroscopy, for further investigations.

Excluded regions : Y-H...X vs. C-H...X

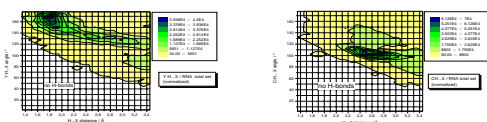


Fig. 2: Contour plots of two-dimensional distributions of the H...X distance and of the Y(C)-H...X angle for standard Y-H...X contacts (Y: O,N) and C-H...X contacts in RNA structure set V (no angle cutoffs)

Two-dimensional distance/angle distributions show an **excluded region** of almost equal shape for standard and C-H...X contacts. For C-H...X contacts the excluded region extends to slightly shorter distances than for standard H-bonds (Fig. 2). In both cases, the **allowed angle range increases with an increasing H...X distance**. The distribution maxima are however completely different:
standard H-bonds: 1.8-2.0 Å, 160-180°;
C-H...X contacts: 2.4-2.8 Å, 100-110°.

The latter maximum corresponds roughly to the van der Waals distance. Nevertheless, there is a substantial number of C-H...X contacts with smaller distances and larger angles.

Interaction types

	r _{H...X} < 3.50 Å	3.00 Å	2.50 Å
total	11225 (100 %)	7686 (100 %)	3138 (100 %)
B-B	8033 (72 %)	6003 (78 %)	2520 (80 %)
B-b	1465 (13 %)	575 (7 %)	191 (6 %)
b-B	990 (9 %)	601 (8 %)	279 (9 %)
b-b	737 (6 %)	507 (7 %)	148 (5 %)

Table 1: Classification of C-H...X interactions according to backbone (B) and base (b) participation; donor parts are given first; r_{H...X}: distance between donor and acceptor (with angle cutoffs)

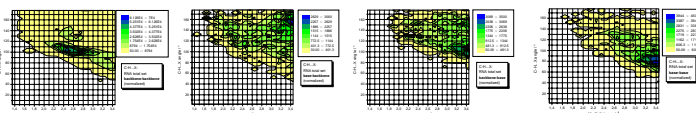


Fig. 3: Contour plots of two-dimensional distributions of the H...X distance and of the C-H...X angle for backbone-backbone, backbone-base, base-backbone and base-base contacts (set V, no angle cutoffs)

In RNA most C-H...X contacts are of the **backbone-backbone** type and occur between **neighbour nucleotides in the sugar phosphate backbone**. Interactions between atoms which are close in atom sequence like **C5'-O3'** and **C2'-O5'** are **fixed** within unfavourable angle ranges. On the other hand, the interactions **C5'-O2'** and **C2'-O4'** show structural **flexibility** and can adopt short distances and large angles (Fig. 4). Backbone-backbone interactions are, however, not restricted to next neighbour contacts (Fig. 5) but have also been found to mediate **tertiary interactions** and to connect nucleotide pairs n/n+2 and n/n+3 (Fig. 5).

O2' is also the most frequently used acceptor in C-H...O **base-backbone** contacts. Very short **C8-O2'** and **C6-O2'** interactions between neighbour nucleotides have been observed in several NMR structures (1rau, 1rng, 1rnk). They are similar to the known intra-residue C8-O5' and C6-O5' interactions. **Backbone-base** C-H...O contacts occur more often than base-backbone interactions, but they tend to be longer than the latter ones (Fig. 3; Table 1).

C2-O2 is the most frequent **base-base** C-H...X motif (Table 2). These interactions are often accompanied by standard hydrogen bonds which may influence their geometry. An especially interesting motif has been found in an RNA tetraplex structure (1rau; Fig. 5). Here **four uracil bases are linked by eight C-H...O bonds**. In several structures, like an aptamer complex (1fmm), RNA-protein complexes (1gr, 1qrs, 1qrt, 1qru), and in the group I ribozyme domain (1gid), **AA base pairs** have been found with a standard hydrogen bond between N6 and N3 and an additional short **C-H...N contact** between C2 and N7 (Fig. 5).

B-B	B-b	b-B	b-b
C2' - O4' (40 %)	C1' - O6 (10 %)	C8 - O2' (19 %)	C2 - O2 (38 %)
C5' - O2' (22 %)	C2' - N7 (9 %)	C6 - O2' (15 %)	C6 - O2 (9 %)
C5' - O3' (9 %)	C1' - O2 (9 %)	C6 - O3' (13 %)	C8 - O4 (7 %)
C2' - O5' (6 %)	C2' - O2 (8 %)	C8 - O3' (11 %)	C5 - O2 (5 %)

Table 2: Most frequent C-H...X motifs (H...X distance < 2.5 Å, with angle cutoffs)

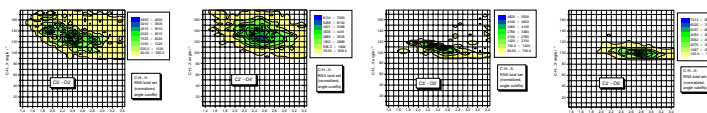


Fig. 4: Contour plots of two-dimensional distributions of the H...X distance and of the C-H...X angle C-H...X for backbone-backbone contacts (set V, with angle cutoffs)

C-H...X structure motifs

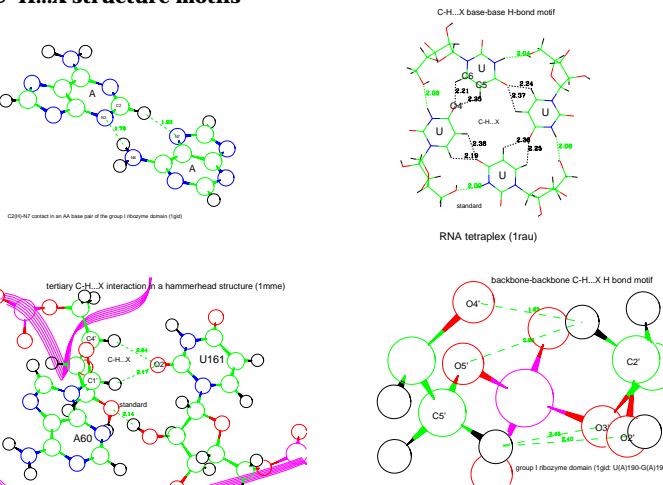


Fig. 5: Examples of C-H...X structure motifs

Conclusion

C-H...X interactions in RNA can be regarded as weak hydrogen bonds. O2' is involved in various frequently occurring C-H...X motifs. As O2' does not occur in deoxyribose these interactions may contribute to structural differences between DNA and RNA.