COSMOLOGICAL IMPLICATIONS OF THE UNCERTAINTY IN H⁻ DESTRUCTION RATE COEFFICIENTS

S. C. GLOVER,¹ ² D. W. SAVIN,³ AND A.-K. JAPPSEN²

Received 2005 June 9; accepted 2005 November 29

ABSTRACT

In primordial gas, molecular hydrogen forms primarily through associative detachment of H⁻ and H, thereby destroying the H⁻ ion. The H⁻ anion can also be destroyed by a number of other reactions, most notably by mutual neutralization with protons. However, neither the associative detachment nor the mutual neutralization rate coefficients are well determined: both may be uncertain by as much as an order of magnitude. This introduces a corresponding uncertainty into the H₂ formation rate, which may have cosmological implications. Here we examine the effect that these uncertainties have on the formation of H₂ and the cooling of protogalactic gas in a variety of situations. We show that the effect is particularly large for protogalaxies forming in previously ionized regions, affecting our predictions of whether or not a given protogalaxy can cool and condense within a Hubble time, and altering the strength of the ultraviolet background that is required to prevent collapse.

Subject headings: atomic data — atomic processes — galaxies; formation — molecular data — molecular processes — stars: formation

1. INTRODUCTION

It has long been argued that molecular hydrogen, H₂, must play a central role in the cooling of primordial gas in the first protogalaxies. This is simply because it is the only coolant present in significant quantities that remains effective at temperatures below 10⁴ K (Saslaw & Zipoy 1967; Peebles & Dicke 1968; Matsuda et al. 1969; Hirasawa 1969; Lepp & Shull 1984). Subsequent numerical work has only served to confirm this (e.g., Hutchings 1976; Palla et al. 1983; Mac Low & Shull 1986; Shapiro & Kang 1987; Teigmark et al. 1997; Machacek et al. 2001; Abel et al. 2002).

It is now clear that one of the keys to understanding the earliest episodes of star formation in the cosmos is a detailed understanding of the chemistry and thermodynamics of H₂. The chemistry of primordial gas has recently been reviewed by a number of authors (Abel et al. 1997; Galli & Palla 1998; Stancil et al. 1998; Lepp et al. 2002). These authors find that considerable chemical complexity is possible despite the limited number of elements available (essentially only hydrogen, helium, and lithium, plus isotopic counterparts such as deuterium or ³He). But much of this complexity arises from the chemistry of minor coolants such as HD or LiH, or trace molecules such as H₂ or HeH⁺. These molecules do not play a significant role in the chemistry of H₂ (Abel et al. 1997; Glover 2001), and the cooling that they provide is generally unimportant compared to that coming from H₂. It is thus usually unnecessary to model their chemistry (although HD cooling can become important in gas with an atomic hydrogen number density n_H > 10⁶ cm⁻³ and temperature T < 200 K; see Flower & Pineau des Forêts 2001 or Nakamura & Umemura 2002).

The omission of these minor and trace molecules allows substantial simplifications to be made to the chemical network of reactions. In particular, Abel et al. (1997) showed that the formation and destruction of H₂ can be accurately followed over a wide range of temperatures and densities with as few as 28 reactions. In many circumstances, this number can be reduced further. For instance, six of the reactions deal with the ionization and recombination of helium and so play no role in gas in which helium remains neutral. Furthermore, four of the photochemical reactions—the photoionization of H, H₂, and H₂⁺, and the photodissociation of H₂ by absorption into the continuum of the Lyman and Werner band systems (Allison & Dalgarno 1969)—require photons with energies above the Lyman limit and so are important only if a strong source of hard ultraviolet photons is present.

The key reactions responsible for the formation of H₂ are easily summarized. Direct radiative association of atomic hydrogen is strongly forbidden (Gould & Salpeter 1963), and so the main gas-phase pathway by which H₂ is formed makes use of the H⁻ ion as an intermediary. H⁻ is formed by the slow radiative association reaction

\[ H + e^- \rightarrow H^- + \gamma, \]  

and is then destroyed by a fast associative detachment reaction with atomic hydrogen, forming H₂:

\[ H^- + H \rightarrow H_2 + e^- . \]  

Molecular hydrogen can also be formed by a similar chain of reactions involving H⁺:

\[ H + H^+ \rightarrow H_2^+ + \gamma, \]  
\[ H + H_2^+ \rightarrow H_2 + H^+ , \]

but in most circumstances the H⁻ pathway dominates (Glover 2003), as we now explain.

Various reactions compete with associative detachment to destroy H⁻, of which the most important are mutual neutralization with protons,

\[ H^- + H^+ \rightarrow H + H, \]

and photodetachment by infrared and/or optical photons,

\[ H^- + \gamma \rightarrow H + e^- . \]

¹ Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192.
² Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany.
³ Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027-6601.
At very high redshifts (z > 100), the cosmic microwave background (CMB) temperature is large enough to produce a substantial photodetachment rate (Galli & Palla 1998; Stancil et al. 1998), and so only a few of the H\(^{-}\) ions survive for long enough to form H\(_2\). At these redshifts, H\(_2\) formation is dominated by the slower H\(_2 \)\(^+\) pathway. However, at lower redshifts, photodetachment of H\(^{-}\) by the CMB rapidly becomes unimportant, and at the redshifts of interest in this paper its effect is negligible. The net rate of H\(_2\) formation is therefore determined by two main factors—the fractional ionization of the gas, which controls the rate of the initiating radiative association (reaction [1]), and the fraction of the resulting H\(^{-}\) that is destroyed by associative detachment (reaction [2]). A similar state of affairs exists for formation of H\(_2\) by the H\(_2 \)\(^+\) pathway, with the main alternative destruction mechanism in this case being dissociative recombination:

\[ \text{H}_2^+ + e^- \rightarrow \text{H} + \text{H}. \]  

Once formed, H\(_2\) can be destroyed by charge transfer with H\(^{+}\),

\[ \text{H}_2 + \text{H}^+ \rightarrow \text{H}_2^+ + \text{H}, \]  

by collisional dissociation with free electrons or atomic hydrogen,

\[ \text{H}_2 + e^- \rightarrow \text{H} + \text{H} + e^-, \]  

\[ \text{H}_2 + \text{H} \rightarrow \text{H} + \text{H} + \text{H}, \]  

or by photodissociation via the Solomon process (Stecher & Williams 1967),

\[ \text{H}_2 + \gamma \rightarrow \text{H} + \text{H}, \]  

although if a large amount of H\(_2\) is present then the molecular gas will "self-shield" as the Lyman and Werner bands become optically thick, reducing the effectiveness of photodissociation.

In addition to the reactions listed above, several others are required to complete our chemical network. Collisional ionization of H by electrons,

\[ \text{H} + e^- \rightarrow \text{H}^+ + e^- + e^-, \]  

and radiative recombination of H\(^{+}\),

\[ \text{H}^+ + e^- \rightarrow \text{H} + \gamma, \]  

must be included if we wish to model the evolution of the fractional ionization of the gas correctly. The other reactions include the dissociative recombination of H\(_2 \)\(^-\) (reaction [7] above), together with

\[ \text{H}^- + e^- \rightarrow \text{H} + e^- + e^-, \]  

\[ \text{H}^- + \text{H} \rightarrow \text{H} + \text{H} + e^-, \]  

\[ \text{H}_2^- + e^- \rightarrow \text{H} + \text{H}^+, \]  

Reactions (7) and (14)–(17) all play a role in regulating either the H\(^{-}\) or the H\(_2 \)\(^-\) abundance, and so affect H\(_2\) formation. However, it should be noted that in most circumstances, the influence of these reactions on the H\(_2\) formation rate is relatively small. In addition to these reactions, Abel et al. (1997) also include an additional reaction between H\(^{-}\) and H\(_2 \)\(^-\),

\[ \text{H}^- + \text{H}_2^- \rightarrow \text{H}_2 + \text{H}, \]  

but they show that in practice this reaction is never significant, and for this reason it is not included in our model.

There is considerable variation in the accuracy with which the rates and rate coefficients of the reactions in this network are known. Some, such as the radiative association reaction that forms H\(^{-}\) (reaction [1]), have rate coefficients that have been determined to a high level of accuracy (in this case because the rate coefficient can be obtained through the use of the principle of detailed balance and the accurately known H\(^{-}\) photodetachment cross section; see the discussion in Abel et al. 1997 for more details). Others, however, have rate coefficients that are considerably more uncertain, such as the case for the H\(_2\) charge transfer reaction (reaction [8]) discussed by Savin et al. (2004).

In this paper, we are concerned with two reactions for which a large degree of uncertainty exists: the destruction of H\(^{-}\) by associative detachment with H (reaction [2]) and by mutual neutralization with H\(^{+}\) (reaction [5]). In situations where collisional processes dominate the destruction of H\(^{-}\), and the effects of photodetachment (reaction [6]) are unimportant, it is the competition between reactions (2) and (5) that determines what fraction of the H\(^{-}\) formed by reaction (1) goes on to form H\(_2\). Therefore, substantial variations in the rate coefficients of these two reactions can result in major variations in the H\(_2\) formation rate. This in turn may have significant consequences for the cooling of primordial gas in situations of cosmological importance, such as in newly forming protogalaxies in formerly ionized regions, or in the partially ionized interface regions that exist around the cosmological H \(\alpha\) regions created by massive metal-free stars (Ricotti et al. 2001). It is therefore important to determine the extent to which the existing uncertainties in the associative detachment and mutual neutralization rate coefficients affect the conclusions that we can draw regarding the formation of H\(_2\) in primordial gas.

To investigate this, we have performed several sets of simulations of protogalactic collapse in which we have varied the rate coefficients for reactions (2) and (5) over a range of plausible values and have determined how sensitive the outcome is for various initial conditions to the particular collision data chosen. In §2 we review the current state of knowledge regarding the values of the associative detachment and mutual neutralization rate coefficients. In §3 we describe the code used to perform our simulations. The initial conditions used in our simulations are discussed in §4, and the results of the simulations are presented in §5. We conclude with a brief discussion in §6.

2. DATA FOR ASSOCIATIVE DETACHMENT AND MUTUAL NEUTRALIZATION

Here we review the published theoretical and experimental work for the associative detachment reaction (2) and the mutual neutralization reaction (5) at temperatures \(T\) and collision energies \(E\) relevant for cosmology. In specific, we focus on work at \(T \leq 10^4 K\) or \(E \leq 1 eV\). The corresponding published rate
coefficients for reactions (2) and (5) are shown in Figures 1 and 2, respectively.

2.1. Associative Detachment

The only measurement at cosmological temperatures of the associative detachment reaction (2) has been carried out by Schmeltekopf et al. (1967). They used a fast-flowing afterglow system and claimed that their measured rate coefficient is accurate to within a factor of 2. Their result is used by essentially all early universe chemical models (e.g., Abel et al. 1997; Galli & Palla 1998; Lepp et al. 2002).

The measurement of Schmeltekopf et al. (1967) was quickly followed with a semiclassical theoretical calculation by Dalgarno & Browne (1967) using a simple extension on the theory for radiative association. Browne & Dalgarno (1969) subsequently revised their earlier calculations using new potentials for the

---

Fig. 1.—Published experimental and theoretical rate coefficients for the associative detachment reaction \( \text{H}^- + \text{H} \rightarrow \text{H}_2^- \rightarrow \text{H}_2 + e^- \) at temperatures relevant for early universe chemistry.

Fig. 2.—Published theoretical and experimental rate coefficients for the mutual neutralization reaction \( \text{H}^- + \text{H}^+ \rightarrow \text{H} + \text{H} \) at temperatures relevant for early universe chemistry.
intermediate $H_2^-$ molecular anion, taking into account the lowest two electronic states, namely, the attractive $X^+ \Sigma_u^+ 2$ state and the repulsive $B \Sigma_u^+ 2$ state. They presented results for their potentials $(1 + 3)$ and $(2 + 4)$. Both of these results from Browne & Dalgarno are shown in Figure 1. Bieniek & Dalgarno (1979) revisited this reaction 10 years later, treating the problem as one of scattering by a complex potential (i.e., resonant scattering theory). All of these calculations are in reasonable agreement with the laboratory results, to within the quoted experimental factor of 2 uncertainty.

Sakimoto (1989) calculated the rate coefficient for reaction (2), following the quantum methodology of Bieniek & Dalgarno (1979). However, Sakimoto paid attention to lower collision energies than considered by Bieniek & Dalgarno and also used new results for the $X^+ \Sigma_u^+ 2$ potential. The resulting rate coefficient is a factor of almost 4 times larger than the measurement of Schmeltekopf et al. (1967). Sakimoto also presents a classical Langevin limit rate coefficient that falls outside of the experimental uncertainty limits.

Using resonant scattering theory, a local complex potential, and a new $X^+ \Sigma_u^+ 2$ potential energy surface, Launay et al. (1991) calculated the rate coefficient for reaction (2). They present results for their potentials V1 and V2. Each resulting rate coefficient lies within the factor of 2 experimental error bars. However, the energy dependence calculated by Launay et al. differs from that of Browne & Dalgarno (1969) above $\approx 2000$ K. Also, it appears that the results using the potential V2 are incorrect, as is discussed in more detail by Čiček et al. (1998).

The most recent calculation for reaction (2) is that of Čiček et al. (1998). They use a nonlocal resonance model and a new potential for the $X^+ \Sigma_u^+ 2$ state. The resulting rate coefficient lies a factor of $\approx 3$ above the experimental results of Schmeltekopf et al. (1967). Čiček et al. suggest that the source of this discrepancy lies with the experimental results.

Čiček et al. (1998) also present a classical Langevin limit rate coefficient that is in good agreement with the experiment but lies a factor of $\approx 2$ below the Langevin results of Sakimoto (1989). This appears to be due to a simple error on the part of Sakimoto. At the low energies relevant here, the initial $H^- + H$ system can approach with equal probability along either the attractive $X^+ \Sigma_u^+ 2$ or the repulsive $B \Sigma_u^+ 2$ potential energy surfaces. Here only the attractive state can lead to associative detachment. Since only half of the collisions will enter along this state, the Langevin rate coefficient must therefore be multiplied by a factor of 0.5 to take this into account. Sakimoto appears to have left this factor out.

Taking the lower limit for the experimental rate coefficient and the highest calculated value yields almost an order of magnitude spread. This is surprisingly large for such a simple reaction. A partial explanation for the disagreement between the various theoretical calculations may lie in the $H^-_2$ potentials used. The three most recent calculations for reaction (2) use differing potentials for the $X^+ \Sigma_u^+ 2$ state (Sakimoto 1989; Launay et al. 1991; Čiček et al. 1998) and all yield differing results. With theory unable to converge to the same value for the rate coefficient, it appears that the only hope of improving our understanding of reaction (2) lies with carrying out new laboratory measurements.

### 2.2. Mutual Neutralization

The first published theoretical rate coefficient for reaction (5) at cosmological temperatures appears to be the Landau-Zener (LZ) results of Bates & Lewis (1955). Moseley et al. (1970) are the only group to have measured reaction (5) at collision energies below 3 eV. They carried out measurements to energies as low as 0.1 eV. Using their cross section results and extrapolating below 0.1 eV, they produced an experimentally derived rate coefficient (Moseley et al. 1970; Peterson et al. 1971). Their experimental rate coefficient is a factor of $\approx 3$ times larger than the LZ results of Bates & Lewis (1955).

Dalgarno & McCray (1973) write that the rate coefficient for

$$H^- + X^+ \rightarrow H + X$$

is “fairly insensitive to the nature of the ions $X^+$.” They recommend a rate coefficient for all such reactions that is based on the results of Peterson et al. (1971) for reaction (5). Surprisingly, the rate coefficient of Dalgarno & McCray is a factor of $\approx 2$ times smaller than that of Peterson et al. This may be a partial explanation as to why Dalgarno & McCray estimate that their recommended rate coefficient is accurate to “perhaps a factor of 2.” Prasad & Huntress (1980), in their commonly cited compilation of rate coefficients for gas phase chemistry, use the recommended rate coefficient of Dalgarno & McCray (1973). This then made its way into the recommended data of Duley & Williams (1984), and then into Shapiro & Kang (1987).

Fussen & Kubach (1986) calculated reaction (5) using a quantum close-coupling treatment. Their cross section results are in good agreement with the LZ calculations of Bates & Lewis (1955) but a factor of $\approx 3$ times smaller than the measurements of Moseley et al. (1970). Croft et al. (1999) used the cross section calculations of Fussen & Kubach to produce a rate coefficient. Lepp et al. (2002) in their early universe chemical network presented and used a fit based on the results of Croft et al.

Dalgarno & Lepp (1987) present two rate coefficients for reaction (5), a high value and a low value. The higher value is said to be based on the results of Moseley et al. (1970), but as can be seen in Figure 2, lies a factor of $\approx 4$ below their results. The lower value is said to be derived from the experimental results of Szucs et al. (1984) and Peart et al. (1985), which lie close to the theoretical cross section predictions of Bates & Lewis (1955). However, Figure 2 shows that the lower recommended value of Dalgarno & Lepp is a factor of $\approx 3$ times smaller than that of Bates & Lewis (1955). Also, these experimental results were carried out for collision energies in excess of 5 eV, which is of little relevance for early universe chemistry.

The cause for the differences for each rate coefficient of Dalgarno & Lepp (1987) and their stated sources is unclear. Perhaps both represent typographical errors. The lower rate coefficient of Dalgarno & Lepp was adopted by Abel et al. (1997) for their primordial chemistry.

Clearly, there is still a large uncertainty in the rate coefficient for reaction (5). It is true that LZ and quantum close-coupling calculations are in good agreement for $T < 10^4$ K. But the only measurement at the relevant collision energies lies a factor of $\approx 3$ higher. Cross section measurements at collision energies above 3 eV have been carried out by a number of groups (Szucs et al. 1984; Peart et al. 1985; Peart & Hayton 1992), and these are in reasonable agreement with the calculations of Bates & Lewis (1955) and Fussen & Kubach (1986). However, since the work of Moseley et al. (1970) no further measurements have been carried out at collision energies relevant for early universe chemistry. In addition, no explanation has been given to date as to why the data of Moseley et al. (1970) may be wrong (cf. Szucs et al. 1984; Peart et al. 1985; Peart & Hayton 1992). It is clear that in order to resolve this issue, new laboratory
3. COMPUTATIONAL METHOD

To aid us in assessing the impact of the existing uncertainties in the associative detachment and mutual neutralization rate coefficients on the cooling and collapse of primordial gas, we have performed a number of simulations of the collapse of gas into protogalactic halos using a numerical technique known as smoothed particle hydrodynamics (SPH; Gingold & Monaghan 1977; Lucy 1977; Monaghan 1992). SPH is a Lagrangian method on both the density and temperature of the gas, as well as on its

where

represents averages over the local fluid flow. For each SPH particle, we therefore must solve a set of chemical equations of the form

\[ \frac{dx_i}{dt} = C_i - D_i x_i, \]

where \( x_i \) is the fractional abundance of species \( i \), computed with respect to the total number density of hydrogen nuclei (i.e., \( x_i = n_i/n \)), where \( n_i \) is the number density of species \( i \) and \( n \) is the number density of hydrogen nuclei. \( C_i \) and \( D_i \) are terms representing the chemical creation and destruction of species \( i \). The values of the creation and destruction terms generally depend on both the density and temperature of the gas, as well as on its chemical composition.

In our code, we follow the chemistry of six species: \( e^- \), \( H^+ \), \( H \), \( H^0 \), \( H_2 \), and \( H_2^+ \). Two of these species, \( H^- \) and \( H_2^+ \), reach chemical equilibrium on very short timescales. In the case of \( H^- \), one can readily show that equilibrium is reached on a timescale

\[ t_{eq,H^-} \simeq \left( k_{rad} n_H + k_{mn} n_{e^-} \right)^{-1}, \]

where \( n_H \) is the hydrogen atom density, \( n_{e^-} \) is the \( H^- \) density, and \( k_{rad} \) and \( k_{mn} \) are the rate coefficients of the charge transfer and dissociative recombination reactions which are the main processes responsible for removing \( H_2^+ \) from the gas. (For reference, these are reactions [4] and [7] in Table 1). Evaluating this expression, we obtain \( t_{eq,H^-} \simeq 1.6 \times 10^9 n_H^{-1} \) s in the limit that \( n_H \ll n_{e^-} \) or \( t_{eq,H^-} \simeq 10^8 n_{e^-}^{-1} \) s in the limit that \( n_H \gg n_{e^-} \). Again, these timescales are both very much shorter than either the cooling time or the free-fall time of the gas.

Because \( H^- \) and \( H_2^+ \) reach equilibrium so quickly, we do not attempt to follow their chemical equilibrium directly in our simulations. Instead, their abundances are computed only as required, under the assumption that both reach equilibrium instantaneously. This approximation introduces a certain amount of error into the computed \( H_2 \) abundance. However, because the time steps used to evolve the SPH particles in our simulations are long compared to \( t_{eq,H^-} \) and \( t_{eq,H_2^+} \), the size of this error will be negligible compared to that arising from the rate coefficient uncertainties that are the subject of this paper. Having made this approximation, we are left with four nonequilibrium species, and thus in theory four chemical equations. However, since charge conservation implies that

\[ x_{H^+} + x_{H_2^+} = x_{e^-} + x_H, \]

measurements are needed for reaction (5) at collision energies below 1 eV.

No. 2, 2006  UNCERTAINTY IN \( H^- \) DESTRUCTION RATE COEFFICIENTS  557

| Reaction Number | Reaction | Reference |
|-----------------|----------|-----------|
| 1               | \( H + e^- \rightarrow H^- + \gamma \) | Wishart (1979) |
| 2               | \( H^+ + H \rightarrow H_2 + e^- \) | See text |
| 3               | \( H + H^+ \rightarrow H_2^+ + \gamma \) | Ramaker & Peck (1976) |
| 4               | \( H + H_2^+ \rightarrow H_2 + H^+ \) | Karpas et al. (1979) |
| 5               | \( H^- + H_2^+ \rightarrow H + H_2 \) | See text |
| 6               | \( H^- + \gamma \rightarrow H + e^- \) | Wishart (1979) |
| 7               | \( H_2^+ + e^- \rightarrow H + H \) | Schneider et al. (1994) |
| 8               | \( H_2 + H^+ \rightarrow H_2^+ + H \) | Savin et al. (2004) |
| 9               | \( H_2 + e^- \rightarrow H + H + e^- \) | Stibbe & Tennyson (1999) |
| 10              | \( H_2 + H \rightarrow H_2 + H + H \) | Mac Low & Shull (1986) |
| 11              | \( H_2 + \gamma \rightarrow H + H \) | Draine & Bertoldi (1996) |
| 12              | \( H + e^- \rightarrow H^- + e^- + e^- \) | Janey et al. (1987) |
| 13              | \( H^+ + e^- \rightarrow H^- + \gamma \) | Ferland et al. (1992) |
| 14              | \( H^+ + e^- \rightarrow H + e^- + e^- \) | Janey et al. (1987) |
| 15              | \( H^- + H \rightarrow H + H + e^- \) | Janey et al. (1987) |
| 16              | \( H^- + H_2 \rightarrow H_2 + e^- \) | Poularet et al. (1978) |
| 17              | \( H_2^+ + \gamma \rightarrow H + H^+ \) | Dunn (1968) |

\( k_{rad} \) and \( k_{mn} \) are the rate coefficients of the charge transfer and dissociative recombination reactions which are the main processes responsible for removing \( H_2^+ \) from the gas. (For reference, these are reactions [4] and [7] in Table 1). Evaluating this expression, we obtain \( t_{eq,H^-} \simeq 1.6 \times 10^9 n_H^{-1} \) s in the limit that \( n_H \ll n_{e^-} \), or \( t_{eq,H^-} \simeq 10^8 n_{e^-}^{-1} \) s in the limit that \( n_H \gg n_{e^-} \). Again, these timescales are both very much shorter than either the cooling time or the free-fall time of the gas.
while conservation of the total number of hydrogen nuclei implies that
\[ x_{\text{H}} + x_{\text{H}^+} + x_{\text{H}^2} + 2x_{\text{H}_2^+} + 2x_{\text{H}_2} = 1, \]  \quad (24)
we can reduce the actual number of equations that we must solve to only two. We choose to solve for \( x_{\text{H}_2} \) and \( x_{\text{H}^+} \), obtaining \( x_c \) and \( x_H \) from equations (23) and (24), but other choices are of course possible. Our choice here should make no difference to our final results.

To solve the chemical equations for a given SPH particle, we make use of a technique known as operator splitting. We assume that within the current time step of the SPH particle, we can solve for the density evolution of the gas separately from its chemical evolution. The density evolution can then be computed using the same prescription as in standard SPH (see Springel et al. 2001 for details), and the updated gas density is then available for use in the chemical equations. These are then solved implicitly using DVODE, a freely available and well-tested double precision ordinary differential equation solver (Brown et al. 1989). Operator splitting introduces a certain amount of error, as in practice the density should vary during the chemical time step. However, the SPH algorithm naturally limits the extent to which the density can change during a single SPH particle time step, by making particles take shorter time steps in rapidly evolving regions. We therefore expect the error introduced by this technique to be small, as is discussed further in § 4.

In common with other authors, we use a simplified reaction network that does not include the chemistry of minor coolants such as HD or LiH. We also neglect any effects due to helium chemistry. Neglect of the minor coolants is justified by the fact that \( \text{H}_2 \) dominates the cooling of the gas for the all of the temperatures and densities found in our simulations (Flower & Pineau des Forêts 2001). At the worst, we may overestimate the final temperature of the gas slightly. Neglect of the helium chemistry is also easily justified, provided that we assume that the bulk of the helium in the gas is in neutral form, as in this case the only reactions involving helium that play any role in determining the \( \text{H}_2 \) abundance—the collisional detachment of \( \text{H}^- \) and collisional dissociation of \( \text{H}_2 \) by \( \text{He} \)—are far less effective than the corresponding reactions with \( \text{H} \) (Abel et al. 1997), so the error in the final \( \text{H}_2 \) abundance will be small. If significant amounts of ionized helium are present, then our assumption introduces a larger error, since we will underestimate the actual electron abundance in the gas, and hence the \( \text{H}^- \) formation rate. However, even in this case, we would expect the error in the final \( \text{H}_2 \) abundance to be relatively small, owing to the small abundance of helium relative to hydrogen. The chemical reactions included in our network are summarized in Table 1. In most cases, we also list the source used for the adopted rate coefficient. The exceptions are the associative detachment and mutual neutralization reactions, which we discuss further below.

Table 1 lists three photochemical reactions: the photo-detachment of \( \text{H}^- \) and the photodissociation of \( \text{H}_2^+ \) and \( \text{H}_2 \). To compute the rates for these reactions, we assumed that our simulated protogalaxies were illuminated by an external background radiation spectrum with the shape of a \( 10^5 \) K blackbody, as should be typical of the brightest Population III stars (Cojazzi et al. 2000). We cut off this spectrum at energies greater than 13.6 eV to account for absorption by neutral hydrogen in the protogalactic gas and in the intergalactic medium. The strength of the background is specified in terms of the flux at the Lyman limit, \( J(\nu_\alpha) = 10^{-21} J_{21} \text{ ergs} \text{s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1} \).

If sufficient \( \text{H}_2 \) forms within the protogalaxy, it will begin to self-shield, reducing the effective photodissociation rate. An exact treatment of the effects of self-shielding is computationally infeasible, as it would require us to solve for the full spatial, angular, and frequency dependence of the radiation field at every time step. Instead, we have chosen to incorporate it in an approximate manner. We assume that the dominant contribution to the self-shielding at a given point in the protogalaxy comes from gas close to that point, and so only include the contribution to the self-shielding that comes from the nearby \( \text{H}_2 \). To implement this approximation numerically, we make use of the fact that Gadget already defines a suitable local length scale: the SPH smoothing length \( h \), which characterizes the scale over which the flow variables are averaged. In Gadget, as in all modern SPH codes, \( h \) is allowed to vary from point to point within the flow and is automatically adjusted in order to keep the mass enclosed within a sphere of radius \( h \) roughly constant. Further details can be found in Springel et al. (2001).

In our calculation of the \( \text{H}_2 \) column density used to compute the degree of self-shielding we include only \( \text{H}_2 \) that lies within a single smoothing length of the point of interest. We justify this approximation by noting that in our simulations, widely separated SPH particles typically move with a significant velocity relative to one another. Consequently, the contribution to the total absorption arising from one particle is Doppler-shifted when viewed from the rest frame of the other particle. If this Doppler shift is large compared to the line widths of the Lyman-Werner band transitions, then the effect is to dramatically reduce the extent to which the absorption contributes to the total self-shielding. On the other hand, gas close to the point of interest will typically have a much smaller relative velocity, and so will contribute far more effectively. Our approximation considers only the latter contribution, and assumes that the former contribution is completely negligible. In practice, of course, the distant gas is likely to contribute to some nonnegligible extent, and so we will tend to underestimate the true amount of self-shielding, unless the gas infall is highly supersonic. Nevertheless, we believe that our approximation remains useful as it is computationally efficient, and also represents an improvement over most previous treatments, which typically assumed that the gas was optically thin (e.g., Machacek et al. 2001; Ricotti et al. 2002). An exception is the work of Yoshida et al. (2003), who include the effects of self-shielding by computing the effective \( \text{H}_2 \) column density seen by each SPH particle in a more detailed fashion than we do here. However, their approach appears to be less computationally efficient, and also will tend to overestimate the amount of self-shielding, as they do not take the effects of velocity into account.

Finally, we assume that ionization from X-rays or cosmic rays is negligible. Previous work suggests that even if a low level of ionization is present, it will not have a major effect on the outcome of the collapse (Glover & Brand 2003; Machacek et al. 2003).

A second major modification that we have made to the Gadget code is the inclusion of a treatment of radiative heating and cooling. Cooling in our model comes from three main sources: electron impact excitation of atomic hydrogen (Ly\( \alpha \) cooling), which is effective only above about 8000 K; rotational and vibrational excitation of \( \text{H}_2 \); and Compton cooling. Rates for Ly\( \alpha \) cooling and Compton cooling were taken from Cen (1992), while for \( \text{H}_2 \) rovibrational cooling we used a cooling function from Le Bourlot et al. (1999). The latter is based in part on collisional excitation rates computed using the Partridge et al. (1993) potential energy surface for the H-H\(_2\) system, which is less
accurate than the more recent Boothroyd et al. (1996) surface. However, comparison between the Le Bourlot et al. (1999) cooling rate and rates computed using the Boothroyd et al. (1996) potential show a reasonable level of agreement (Flower et al. 2000). In any case, since we use the same H2 cooling function in all of our runs, any uncertainty present in it is unlikely to affect our conclusions. In models where an ultraviolet background is present, we include the effects of heating from the photodissociation of H2, assuming that 0.4 eV of energy per photodissociation is deposited as heat (Black & Dalgarno 1977). We also include heating due to the ultraviolet pumping of H2, following Burton et al. (1990), although this is only important in high-density gas ($n \geq 10^4$ cm$^{-3}$).

To incorporate the radiative heating and cooling terms into Gadget, we again use an operator splitting technique. In this case, we assume that the change in the internal energy of the gas due to pressure work can be computed separately from that due to radiative heating and cooling. The former can then be calculated in the same fashion as in the standard Gadget code, while the latter can be computed by solving

$$\frac{d\epsilon}{dt} = \Gamma - \Lambda,$$

(25)

where $\epsilon$ is our initial estimate of the internal energy density of the gas, which already includes the effects of the pressure work term, $\Gamma$ is the radiative heating rate per unit volume, and $\Lambda$ is the radiative cooling rate per unit volume. We solve this equation implicitly using DVODE at the same time that we solve the chemical equations. Again, the use of operator splitting introduces some error into the thermal evolution of the gas, but, as before, we expect this error to be small (see § 4 for further discussion).

Finally, to allow us to represent gas that has collapsed beyond the resolution limit of the simulation in a numerically robust manner, we have modified the code to allow it to create sink particles—massive, non-gaseous particles, designed to represent dense cores, that can accrete gas from their surroundings but otherwise interact only via gravity (Bate et al. 1995). The design and implementation of our sink particle algorithm is discussed elsewhere (Jappsen et al. 2005).

4. INITIAL CONDITIONS

As we wish to be able to run a large number of simulations of protogalactic collapse, we have chosen to limit the computational cost of each simulation by starting from somewhat simplified initial conditions. Since we are not particularly interested in the assembly history of the dark matter halo in which the protogalaxy sits, or in studying the response of the halo to the cooling of the gas, we chose to model the influence of the halo by using a fixed background potential. To construct this potential, we assumed that the halo was spherically symmetric, with the density profile of a Burkert halo (Burkert 1995):

$$\rho_{\text{dm}}(r) = \frac{\rho_{\text{dm},0}}{1 + (r^2 / r_c^2)},$$

(26)

We took the central density of the halo to be $\rho_{\text{dm},0} = 0.13 M_\odot$ pc$^{-3}$ and also specified the total mass of dark matter in the halo, $M_{\text{dm}} = 10^7 M_\odot$. The core radius $r_c$ is calculated internally by Gadget, based on $M_{\text{dm}}$ and the initial redshift $z$. For the gas, we assumed an initially uniform distribution, with an initial density $\rho_g$, taken to be equal to the cosmological background density. The initial temperature of the gas was also taken to be uniform, with a value $T_g$, the choice of which is discussed below.

The quantity of gas present in our simulations was taken to be a fraction $\Omega_h/\Omega_{\text{dm}}$ of the total mass of dark matter, where $\Omega_{\text{dm}} = \Omega_m - \Omega_h$. We took values for the cosmological parameters from Spergel et al. (2003), and so $\Omega_h = 0.047$ and $\Omega_m = 0.29$, giving us a total gas mass of $M_g = 0.047(0.29 - 0.047) \times 10^7 M_\odot = 1.934 \times 10^6 M_\odot$. In most of our simulations, we used 32,768 SPH particles to represent this gas, giving each SPH particle a mass $M_{\text{part}} \approx 59 M_\odot$. In order to properly resolve gravitationally bound clumps (or other gravitationally bound structures) in SPH simulations, they must be represented by at least twice as many SPH particles as are used in the SPH smoothing kernel (Bate & Burkert 1997). In our simulations, our smoothing kernel encompasses approximately 40 particles—pros reasons of computational efficiency, the number is allowed to vary slightly, but never by more than five particles—and so our minimum mass resolution is $M_{\text{res}} \approx 80 M_{\text{part}} \approx 4720 M_\odot$. To verify that our results are insensitive to the value of $M_{\text{res}}$, we have repeated several of our simulations using a larger number of SPH particles, 131,072, corresponding to a 4 times smaller value of $M_{\text{res}}$. We find only minor differences (of the order of 10% or less) between the temperature and density evolution in these runs and in the lower resolution runs discussed below.

Since the SPH particle time steps in these higher resolution runs will typically be shorter than those in our lower resolution runs, the size of the error introduced by our use of operator splitting will also be different. Therefore, the close agreement of the results of these runs with their lower resolution counterparts also helps to reassure us that the splitting errors in both sets of runs are small.

To prevent artificial fragmentation or other numerical artifacts from affecting our results, it is necessary either to halt the simulation before the local Jeans mass, $M_J$, falls below $M_{\text{res}}$ in any part of the simulation volume, or to use sink particles to represent regions where $M_J < M_{\text{res}}$. We have chosen the latter course, and so create sinks in regions where the gas density exceeds $6 M_\odot$ pc$^{-3}$ using the algorithm described in Jappsen et al. (2005). This corresponds to a hydrogen atom number density $n_{\text{crit}} \approx 200$ cm$^{-3}$. The Jeans mass is defined here as

$$M_J = \frac{\pi^{5/2}}{6} G^{-3/2} \rho_g^{-1/2} c_s^3,$$

(27)

where $\rho_g$ is the gas density and $c_s$ is the adiabatic sound speed. At this value of $n_{\text{crit}}$, we find $M_J$ has a value of

$$M_J \approx 5000 \left( \frac{T}{100 \text{ K}} \right)^{3/2} M_\odot.$$  

(28)

This is greater than $M_{\text{res}}$ provided that the temperature of the dense gas exceeds 100 K. As we shall see later, the minimum gas temperature reached by dense, gravitationally collapsed gas in our simulations is typically no smaller than 150 K and so our simulations remain well resolved up to the point at which sink particles are created.

Once created, sink particles can accrete gas from their surroundings, provided that the SPH particles representing the gas (1) come within a specified radius of the sink particle (the accretion radius $r_{\text{acc}}$), (2) are gravitationally bound to the sink, and (3) satisfy certain other conditions (see Jappsen et al. 2005 for details). In our simulations, we set $r_{\text{acc}} = 20$ pc. This value

5 Note that in this definition, we assume that the self-gravity of the gas dominates on small scales within the halo, since the gas is dissipative and so can collapse to much higher densities than the nondissipative dark matter.
is chosen to be somewhat larger than Jeans length of the gas at \( n_{\text{crit}} \), but it is admittedly somewhat arbitrary. However, since SPH particles within \( r_{\text{acc}} \) will only be accreted if they satisfy all of the necessary criteria, the outcome of our simulations should be insensitive to the precise value chosen for \( r_{\text{acc}} \).

We initialized each of our simulations at a redshift \( z = 20 \) and allowed them to run for 220 Myr; given our adopted cosmological parameters, this interval corresponds to approximately 1.25 Hubble times, with the simulations terminating at a redshift \( z \approx 11.2 \). Protogalaxies that fail to cool and collapse during this interval are unlikely to get the chance to do so thereafter, as the typical interval between major mergers of dark matter halos is of the order of a Hubble time (Lacey & Cole 1993).

Finally, it should be noted that the initial conditions described here are undoubtedly highly simplified and are unlikely to properly represent the full hydrodynamics of protogalactic formation. However, they should be more than sufficient for our purposes, as we are primarily interested in the difference in the outcomes of simulations run using different values for the mutual neutralization and associative detachment rate coefficients, and in this respect we would expect the behavior of these simple models to be a reliable guide to the behavior of more complex, but ultimately more realistic models.

5. RESULTS

To investigate how sensitive the cooling and collapse of protogalactic gas are to the choice of mutual neutralization and/or associative detachment rate coefficients, we have performed five sets of simulations, with initial conditions as summarized in the previous section and in Table 2. For each set of parameters listed in Table 2, we performed nine separate runs in which the mutual neutralization and associative detachment rate coefficients were varied individually. For runs 1–3 we adopted the lower value for the mutual neutralization rate coefficient from Dalgarno & Lepp (1987), while for the associative detachment rate coefficient we used values of \( 0.65 \times 10^{-9} \), \( 1.30 \times 10^{-9} \), and \( 5.00 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1} \), respectively, where the central value corresponds to the value measured by Schmeltekopf et al. (1967) and the other values represent lower and upper bounds on plausible values, as previously discussed in §2. For runs 4–6 we adopted the mutual neutralization rate coefficient of Croft et al. (1999) and varied the associative detachment data as before. Finally, for runs 7–9 we repeated this procedure using the mutual neutralization rate coefficient of Peterson et al. (1971). For ease of reference, the nine different combinations of rate coefficients used are listed in Table 3.

5.1. Cold Initial Conditions

Our first set of simulations (runs A1–A9) examined collapse beginning from cold initial conditions. In these runs we took initial values for the fractional ionization and \( \text{H}_2 \) abundance from the intergalactic medium (IGM) chemistry model of Stancil et al. (1998), and adopted an initial gas temperature \( T_g = 12 \text{ K} \), corresponding to the value of the IGM temperature at \( z = 20 \) in the absence of any form of preheating.

The state of the gas at the end of run A1 is illustrated by the plot of temperature versus number density shown in Figure 3. Infalling cold gas has been shock-heated to \( T > 10^5 \text{ K} \) within the potential well of the dark matter halo, and some cooling of the highest density gas is evident, although this cooling is not yet far advanced. By the end of the simulation, the density of the central gas has reached \( \tau_c \approx 5 \times 10^{-2} \text{ cm}^{-3} \) and its temperature is \( T_c \approx 6900 \text{ K} \) (although note that because of the nature of the SPH algorithm, these values actually represent a weighted average over gas within a single smoothing length of the center of the halo). Its fractional ionization remains fairly low.

---

**Table 2**

| Run | \( T_g \) (K) | \( x_e \) | \( x_{\text{H}_2} \) | \( J_{\text{H}_2} \) |
|-----|----------------|----------|----------------|----------------|
| A... | 12 | 2.2 \times 10^{-4} | 2.4 \times 10^{-6} | 0.0 |
| B... | 10^4 | 1.0 | 0.0 | 0.0 |
| C... | 10^4 | 1.0 | 0.0 | 10^{-2} |
| D... | 10^4 | 1.0 | 0.0 | 3 \times 10^{-3} |
| E... | 10^4 | 1.0 | 0.0 | 10^{-3} |

**Table 3**

| Run | Mutual Neutralization | Associative Detachment |
|-----|----------------------|------------------------|
| 1... | 7 \times 10^{-7}T^{-1/2} | 0.65 \times 10^{-9} |
| 2... | 7 \times 10^{-7}T^{-1/2} | 1.30 \times 10^{-9} |
| 3... | 7 \times 10^{-7}T^{-1/2} | 5.00 \times 10^{-9} |
| 4... | 2.4 \times 10^{-7}(1 + 5 \times 10^{-5})T^{-1/2} | 0.65 \times 10^{-9} |
| 5... | 2.4 \times 10^{-7}(1 + 5 \times 10^{-5})T^{-1/2} | 1.30 \times 10^{-9} |
| 6... | 2.4 \times 10^{-7}(1 + 5 \times 10^{-5})T^{-1/2} | 5.00 \times 10^{-9} |
| 7... | 5.7 \times 10^{-7}T^{-1/2} + 6.3 \times 10^{-8} | 0.65 \times 10^{-9} |
| 8... | 5.7 \times 10^{-7}T^{-1/2} + 6.3 \times 10^{-8} | 1.30 \times 10^{-9} |
| 9... | 5.7 \times 10^{-7}T^{-1/2} + 6.3 \times 10^{-8} | 5.00 \times 10^{-9} |
| 10... | -9.2 \times 10^{-11}T^{1/2} + 4.4 \times 10^{-11}T | 0.65 \times 10^{-9} |

**Notes.**—We list here all nine possible combinations, together with the number used elsewhere in the paper to refer to each combination. The mutual neutralization rate coefficients are taken from Dalgarno & Lepp (1987), Croft et al. (1999), and Peterson et al. (1971), respectively. For the associative detachment rate coefficients, the central value is taken from Schmeltekopf et al. (1967); the other values represent lower and upper bounds on plausible values.
detachment rate on future H$_2$ formation to be slight. Ultimately, while the differences that we find between the runs are insensitive to the value of the mutual neutralization rate coefficient, we find more evidence for cooling in these runs, in line with the theoretical expectations, but variations in the associative detachment and mutual neutralization rate coefficients continue to have no more than a small effect. For a given set of associative detachment and mutual neutralization data, as we vary $z$ the largest variation in the central gas density is on the order of 5% and the largest variation in the central H$_2$ abundance on the order of 15%. We therefore conclude that our $z = 20$ runs are giving us an accurate picture of the influence of the uncertainties in the rate coefficients in this particular scenario. The fact that the results of these runs are insensitive to the value of the mutual neutralization rate coefficient is easy to understand if one considers that at a fractional ionization of order $10^{-4}$, the net rate of destruction of H$^-$ by mutual neutralization is 2–4 orders of magnitude smaller than the rate of destruction by associative detachment, meaning that the latter always dominates. The sensitivity of the results to the choice of associative detachment rate coefficient comes about in this case due to the influence of a third process, the collisional dissociation of H$^-$ by atomic hydrogen (reaction [15]), which has a rate coefficient which is competitive with that of associative detachment at high temperature. However, by the end of the simulations, the collisional dissociation rate in the central gas has already become smaller than the associative detachment rate, and so we would expect the influence of the uncertainty in the associative detachment rate on future H$_2$ formation to be slight.

The fact that gas in these runs fails to cool effectively is not entirely unexpected, given our choice of halo mass and initial redshift. Previous work on protogalaxy formation in the cold IGM by Tegmark et al. (1997) and others suggests that only gas in halos more massive than some critical mass $M_{\text{crit}}$ will cool effectively (reviewed recently by Bromm & Larson [2004], Ciardi & Ferrara [2005], and Glover [2005]). For gas collapsing at $z < 20$, Tegmark et al. (1997) find $M_{\text{crit}} \gtrsim 10^7 M_\odot$, suggesting that cooling in our $10^7 M_\odot$ halos will only be marginally effective. We therefore performed two additional sets of simulations, using similar initial conditions to runs A1–A9, but starting at $z = 30$ and $z = 40$. Appropriate adjustments were of course made to the initial density and temperature of the cold gas. We find more evidence for cooling in these runs, in line with the theoretical expectations, but variations in the associative detachment and mutual neutralization rate coefficients continue to have no more than a small effect. For a given set of associative detachment and mutual neutralization data, as we vary $z$ the largest variation in the central gas density is on the order of 5% and the largest variation in the central H$_2$ abundance on the order of 15%. We therefore conclude that our $z = 20$ runs are giving us an accurate picture of the influence of the uncertainties in the rate coefficients in this particular scenario.

### 5.2. Hot Initial Conditions

Since the uncertainties in the rate coefficients do not appear to have a large impact on the outcome of protogalactic collapse that begins from cold initial conditions, we chose in our next set of runs (B1–B9) to examine an alternative situation, in which we might expect the uncertainties to have a greater effect. We took as initial conditions gas that was hot ($T = 10^5$ K) and fully ionized ($x_{e^-} = 1.0$, $x_{H_1} = 0.0$). The physical situation that these initial conditions are intended to represent is that of a protogalaxy forming within what Oh & Haiman (2003) term a “fossil” H$^+$ region—an H$^+$ region surrounding an ionizing source that has just switched off, but where the surrounding gas has not yet had time to cool and recombine. H$_2$ formation via associative detachment should be quite efficient in such partially ionized regions. This was first pointed out by Ricotti et al. (2001). Furthermore, prior to cosmological reionization, these regions should be relatively common, since the characteristic lifetimes of the likely ionizing sources—massive Population III stars and/or active galactic nuclei—are significantly shorter than the Hubble time. Indeed, many such transient H$^+$ regions are seen in the simulations of Ricotti et al. (2002).

In Figure 4 we show the state of the gas at the end of run B1. The asterisk on the right-hand side of the plot represents a sink particle of mass $M_{\text{sink}} \sim 1.2 \times 10^5 M_\odot$, which is located at the center of the protogalaxy. Since we have no information on the true temperature and density distribution of the gas represented by the sink particle, beyond the knowledge that its density is greater than 200 cm$^{-3}$, we cannot assign the material within the sink particle to its proper location in this plot. We have therefore assigned it a nominal density of 200 cm$^{-3}$ and temperature of 200 K, as these values are broadly representative of the properties of the gas immediately prior to sink particle formation. The contrast between this figure and Figure 3 is striking. Far more high-density gas is present, especially if one takes the material in the sink particle into account, and the temperature of the gas is much lower: the minimum temperature reached by the end of the simulation in gas that has not become part of the sink is 347 K, while the lowest temperature reached in any of the gas during the course of the run is 172 K. Note that the reason that the latter value is lower is that we cease tracking the temperature of gas once it becomes part of a sink particle. Consequently, the minimum temperature appears to increase after the creation of the sink particle. Compared to these values, the

| Run     | $T_c/10^3$ (K) | $n_e/10^{-2}$ (cm$^{-3}$) | $x_{e^-}/10^{-4}$ | $x_{H_2}/10^{-4}$ |
|---------|----------------|--------------------------|-------------------|------------------|
| A1      | 6.94           | 6.47                     | 1.98              | 1.83             |
| A2      | 6.66           | 6.75                     | 1.98              | 3.08             |
| A3      | 6.30           | 7.30                     | 1.98              | 5.46             |
| A4      | 7.72           | 6.54                     | 1.98              | 0.66             |
| A5      | 7.44           | 6.71                     | 1.98              | 1.24             |
| A6      | 6.35           | 7.32                     | 1.98              | 5.06             |
| A7      | 6.92           | 6.48                     | 1.98              | 1.85             |
| A8      | 6.66           | 6.79                     | 1.96              | 3.00             |
| A9      | 6.30           | 7.25                     | 1.97              | 5.17             |
minimum temperature reached in run A1, which is 6940 K, is a factor of 20–40 times larger.

The reason for this dramatic difference becomes clear if one compares the growth of the mass-weighted mean H$_2$ fraction in runs A1 and B1. This quantity is given by

$$x_{\text{H}_2; \text{mean}} = \frac{\sum_i x_{\text{H}_2; i}}{N_{\text{sph}}}.$$  

where $x_{\text{H}_2; i}$ is the fractional H$_2$ abundance of the $i$th SPH particle and we sum over all $N_{\text{sph}}$ SPH particles. Since all of the SPH particles have the same mass, this gives us a mass-weighted value. The evolution of $x_{\text{H}_2; \text{mean}}$ during the two runs is plotted in Figure 5. Although run A1 has a higher initial H$_2$ abundance than run B1, it is clear that H$_2$ forms far more rapidly in the latter run, due to the high initial fractional ionization, and it very quickly overtakes run A1. Similar differences are seen if instead of comparing the mean H$_2$ fraction, we compare the evolution of the H$_2$ fraction at the center of the density distribution, $x_{\text{H}_2; c}$. Note also that the apparent decline in $x_{\text{H}_2; \text{mean}}$ at $t = 200$ Myr is due to the formation of a sink particle, as explained in the text.

Figure 6 demonstrates that the differences in the associative detachment and mutual neutralization rate coefficients used in these runs lead to large differences in the resulting H$_2$ abundances. For instance, the difference between the value of $x_{\text{H}_2; c}$ in runs B3 and B7 at $t = 150$ Myr is greater than an order of magnitude; and while there is some indication that the value of $x_{\text{H}_2; c}$ in the different runs begins to converge at late times, the differences between the runs remain significant at the end of the simulations. It should also be noted that the sense of the differences is precisely what we would expect on the basis of our previous discussion of the chemistry: H$_2$ formation is more
efficient in runs with a large value of the associative detachment rate coefficient than in runs with a smaller value, but less effective when the mutual neutralization rate coefficient is large than when it is small.

Figures 7 and 8 follow the evolution of the density and temperature at the center of the protogalaxy and demonstrate the effect that these differences in H$_2$ abundance have on the thermal and dynamical evolution of the gas. For the first 70 Myr there is very little difference between the runs. In each case, the gas initially cools to $T \approx 4000$ K through a combination of Ly$\alpha$ cooling (at $T > 8000$ K) and Compton cooling, before gradually being reheated by adiabatic compression in the gravitational potential well of the dark matter. At later times, the gas begins to cool more strongly through H$_2$ rovibrational emission, and the ensuing loss of pressure support leads to the gas rapidly collapsing to high densities. However, the time of onset of this phase of cooling and collapse is sensitive to the H$_2$ abundance of the gas, and so occurs significantly later in runs with small H$_2$ abundances than in runs with high H$_2$ abundances. The physical reason for this dependence is clear: since H$_2$ is the dominant coolant in the gas at $t \approx 70$ Myr, the cooling time of the gas scales roughly as $t_{\text{cool}} \propto x_{\text{H}_2}^{-1}$, and so protogalaxies in which the gas forms less H$_2$ naturally take longer to cool.

One important consequence of this is that the runs display a wide variation in the amount of cold dense gas available for star formation. This is illustrated by our plot of $f_{\text{cc}}$, the fraction of cool, condensed gas, in Figure 9, which shows that the fraction of cold, dense gas varies from 0% (in runs B4, B7, and B8) to almost 25% (in run B3). Given a sufficiently long time to evolve, it is likely that a significant amount of cold gas will eventually accumulate in the runs with less efficient H$_2$ formation. However, we must recall that unlike our idealized model protogalaxies, real protogalaxies are not isolated systems, and do not have an unlimited time in which to evolve before being disrupted by mergers or other external events (such as nearby galactic outflows; see Thacker et al. 2002). Therefore, our determination of whether or not a given halo can cool quickly enough to form stars may depend on our choice of associative detachment and mutual neutralization rate coefficients.

5.3. Cooling and Collapse with a UV Background

So far, we have assumed that the effects of any external UV background are negligible. However, this probably is not the case, as it is likely that a strong UV background will already be present by $z = 20$ (Haiman et al. 2000; Glover & Brand 2003). Indeed, if the Wilkinson Microwave Anisotropy Probe (WMAP) polarization results are taken at face value, they imply that the optical depth to Thomson scattering seen by the CMB is $\tau = 0.17 \pm 0.04$ (Spergel et al. 2003), which means that cosmological reionization probably occurs somewhere within the redshift range $11 < z < 30$ (Kogut et al. 2003) and which therefore suggests that a strong UV background must already be in place at $z = 20$. However, it should be noted that a more recent analysis by Tegmark et al. (2004) that uses data from both WMAP and the Sloan Digital Sky Survey finds a smaller value.
for the optical depth, \( \tau = 0.12^{+0.08}_{-0.06} \), which may be consistent with reionization at a much lower redshift of \( z \sim 6 - 7 \), and hence with there being a much weaker UV background at \( z = 20 \).

To explore how the presence of a UV background may influence our conclusions, we have run several sets of simulations using the same initial conditions as runs B1–B9, but in which the strength of the UV background was varied as listed in Table 2.

In Figures 10–12 we show how the central \( \mathrm{H}_2 \) abundance varies with time in runs C1–C9, D1–D9, and E1–E9, which correspond to UV field strengths specified by \( J_{21} = 10^{-2}, 3 \times 10^{-3}, \) and \( 10^{-3} \), respectively. From Figure 10 it is apparent that the presence of a strong UV background significantly enlarges the difference in outcomes between runs with different associative detachment and mutual neutralization rate coefficients. For instance, the central \( \mathrm{H}_2 \) abundance at the end of run C3 differs from that at the end of run C7 by almost a factor of 100, whereas runs B3 and B7 differ in this respect by no more than a factor of 4. However, from Figures 11 and 12 we see that as the field weakens, the differences become less pronounced; indeed, the results of runs E1–E9 are close to those of runs B1–B9.

The reason for the large spread in outcomes that we see in runs C1–C9 becomes clearer once we consider that photodissociation by the imposed UV background will set an upper limit on the \( \mathrm{H}_2 \) abundance in optically thin gas, given by the equilibrium value at which formation balances photodissociation:

\[
x_{\mathrm{H}_2, eq} \simeq \frac{R_{\mathrm{H}_2}}{R_{\text{pd}}} x_{\mathrm{H}^-}.
\]

Here \( R_{\text{pd}} \) is the photodissociation rate per \( \mathrm{H}_2 \) molecule, and \( R_{\mathrm{H}_2} \) is the rate of \( \mathrm{H}_2 \) formation per H atom. To compute the latter, we note that the dominant contribution comes from \( \mathrm{H}_2 \) formation via the \( \mathrm{H}^- \) pathway, and so we can approximate \( R_{\mathrm{H}_2} \) by the product of the \( \mathrm{H}^- \) formation rate per H atom and the fraction of \( \mathrm{H}^- \) that successfully forms \( \mathrm{H}_2 \) (rather than being destroyed by mutual neutralization). Therefore,

\[
R_{\mathrm{H}_2} = k_{\mathrm{H}^-} n_e \frac{k_{\text{ad}} n_{\mathrm{H}^+}}{k_{\text{ad}} n_{\mathrm{H}^+} + k_{\text{mn}} n_{\mathrm{H}^+}},
\]

where \( k_{\mathrm{H}^-} \) is the rate coefficient for the formation of \( \mathrm{H}^- \) by radiative association. In gas with a high fractional ionization, \( k_{\text{mn}} n_{\mathrm{H}^+} \gg k_{\text{ad}} n_{\mathrm{H}^+} \), and so this expression reduces to

\[
R_{\mathrm{H}_2} \simeq k_{\mathrm{H}^-} \frac{k_{\text{ad}} n_e}{k_{\text{mn}} n_{\mathrm{H}^+}},
\]

which can be further simplified to

\[
R_{\mathrm{H}_2} \simeq k_{\mathrm{H}^-} \frac{k_{\text{ad}}}{k_{\text{mn}}} n_{\mathrm{H}^+},
\]

provided that \( n_e \approx n_{\mathrm{H}^+} \). After 100 Myr of evolution, the gas at the center of our simulated protogalaxies in runs C1–C9 has a temperature of \( T \simeq 7100 \) K and an atomic hydrogen number density of \( n_{\mathrm{H}} \approx 2.6 \times 10^{-2} \) cm\(^{-3}\); the variation in these values between the different runs is about 5%–10%. Using these values
for $n_H$ and $T$, we find that $k_{ff} = 3 \times 10^{-15}$ cm$^3$ s$^{-1}$ (Wishart 1979), and hence that $R_{ff} = 7.8 \times 10^{-17} (k_{ad}/k_{mm})$ s$^{-1}$. Now, since $R_{ad} = 1.3 \times 10^{-14}$ s$^{-1}$, given our assumed UV field strength of $J_{21} = 10^{-2}$ (Draine & Bertoldi 1996), it is easy to show that

$$x_{H_2, eq} \simeq 6.0 \times 10^{-3} \left( \frac{k_{ad}}{k_{mm}} \right),$$

(34)

where the size of the term in brackets varies from $5 \times 10^{-3}$ to $6 \times 10^{-1}$ depending on the values chosen for the rate coefficients. The equilibrium $H_2$ abundance in optically thin gas therefore varies from $3.1 \times 10^{-3}$ to $3.6 \times 10^{-3}$, depending on our choice of chemical rate coefficients. If we take the effects of self-shielding into account, we will obtain slightly larger values, but if anything the spread in outcomes will be even greater, since gas with a large $H_2$ abundance can self-shield far more effectively than gas with a small $H_2$ abundance.

If we now compute the cooling timescale for this gas, which is given by

$$t_{cool} \simeq \frac{3/2 n_{tot} kT}{\Lambda},$$

(35)

where $n_{tot}$ is the total number density of particles (including electrons), and where $\Lambda$ is the cooling rate, in units of ergs cm$^{-3}$ s$^{-1}$, computed using the SPH code for a range of different $H_2$ abundances (see Fig. 13), then we can determine how much $H_2$ is required in order to cool the gas before the end of the simulation. From Figure 13 we see that an $H_2$ abundance $x_{H_2} \gtrsim 10^{-4}$ is required (as here we are at 100 Myr and the simulations run for 220 Myr). Therefore, for those of our runs in which $x_{H_2, eq}$ is larger than this value, we would expect the gas to cool efficiently, while in runs in which $x_{H_2, eq}$ is not sufficiently large, we would expect the gas to cool slowly, if at all. If we examine the evolution of the temperature of the gas at the center of the protogalaxies simulated in these runs, which we plot in Figure 14, we find that our results are consistent with these expectations: in runs C6, C8, and C9, where $x_{H_2, eq} < 10^{-4}$, very little cooling occurs, while in the remainder of the runs, which have $x_{H_2, eq} > 10^{-4}$, far more cooling occurs. It is also clear from the plot that the final temperature reached in the runs that cool depends strongly on the value of $x_{H_2, eq}$. The differences between the thermal evolution of the gas in the various runs are also reflected in the dynamical evolution, as can be seen from Figure 15. This example therefore gives a particularly striking demonstration of the need for better rate coefficient data, since without them we cannot determine which of the various outcomes illustrated in Figures 10, 14, and 15 is actually the correct one.

It is also interesting to compare the length of the cooling time in these runs with the free-fall time, given by

$$t_{ff} = \sqrt{\frac{3\pi}{32 G \rho_{tot}}},$$

(36)
where \( \rho_{\text{tot}} \) is the total matter density (i.e., the sum of the gas density and the dark matter density). At this point of 100 Myr in the simulation, the main contribution to \( \rho_{\text{tot}} \) comes from the dark matter. Evaluating \( t_{\text{ff}} \) at the center of the dark matter halo, we find that it has a value of approximately 20 Myr. From Figure 13 it is clear that for the gas to have \( t_{\text{cool}} < t_{\text{ff}} \), it must have an \( H_2 \) abundance \( x_{H_2} \gtrsim 10^{-3} \). Since this value is an order of magnitude larger than the abundance required to allow cooling by the end of the simulation, i.e., within what is essentially a Hubble time \( t_{\text{H}} \), it is possible for there to be protogalaxies in which \( t_{\text{ff}} < t_{\text{cool}} < t_{\text{H}} \). In these protogalaxies, the timescale on which the gas collapses will be determined by \( t_{\text{cool}} \) rather than \( t_{\text{ff}} \), and so will be directly sensitive to the amount of \( H_2 \) within the gas. Therefore, predictions that we make concerning the dynamical evolution of such protogalaxies will be highly sensitive to our choices for the mutual neutralization and associative detachment rate coefficients, as can already been seen from Figure 15.

![Figure 16](image1)

Fig. 16.—Same as Fig. 8, but for runs D1–D9.

6. DISCUSSION

The results that we have presented in this paper serve as a demonstration of the potential impact that the existing uncertainties in the values of the associative detachment and mutual neutralization rate coefficients, and to the strength of the UV background. In Table 5 we list the value of \( f_{\text{cc}} \) at the end of the simulation for all of the runs in sets B, C, D, and E. The trend of increasing \( f_{\text{cc}} \) with decreasing field strength is clear and is easily understood given our discussion above. However, the table also demonstrates the degree of variation in \( f_{\text{cc}} \) that is introduced by the uncertainties in the rate coefficients. In particular, we see that the lack of reliability in the atomic data introduce an uncertainty of almost an order of magnitude into the value of the UV field strength that is required to prevent gas cooling and collapse.

![Figure 17](image2)

Fig. 17.—Same as Fig. 8, but for runs E1–E9.

If the strength of the UV background is reduced, as in runs D1–D9 or E1–E9, then the value of \( x_{H_2,eq} \) increases, and so the uncertainty in the rate coefficients has less influence, as it is easier for the gas to form enough \( H_2 \) to provide efficient cooling. Nevertheless, the differences between the runs remain significant, as can be seen from the plots of \( H_2 \) abundance versus time in Figures 11 and 12, or the plots of temperature versus time shown in Figures 16 and 17.

One final comparison that we can make between the runs is to examine how sensitive \( f_{\text{cc}} \) is to the values of the associative detachment and mutual neutralization rate coefficients, and to the strength of the UV background. In Table 5 we list the value of \( f_{\text{cc}} \) at the end of the simulation for all of the runs in sets B, C, D, and E. The trend of increasing \( f_{\text{cc}} \) with decreasing field strength is clear and is easily understood given our discussion above. However, the table also demonstrates the degree of variation in \( f_{\text{cc}} \) that is introduced by the uncertainties in the rate coefficients. In particular, we see that the lack of reliability in the atomic data introduce an uncertainty of almost an order of magnitude into the value of the UV field strength that is required to prevent gas cooling and collapse.

### Table 5

| Run | \( J_{21} = 10^{-2} \) | \( J_{21} = 3 \times 10^{-3} \) | \( J_{21} = 10^{-3} \) | \( J_{21} = 0.0 \) |
|-----|-----------------|-----------------|-----------------|-----------------|
| 1   | 0.00            | 0.00            | 0.08            | 0.12            |
| 2   | 0.00            | 0.08            | 0.14            | 0.17            |
| 3   | 2 \times 10^{-4}| 0.15            | 0.19            | 0.23            |
| 4   | 0.00            | 0.00            | 0.00            | 0.00            |
| 5   | 0.00            | 0.00            | 0.005           | 0.08            |
| 6   | 0.00            | 0.10            | 0.15            | 0.18            |
| 7   | 0.00            | 0.00            | 0.00            | 0.00            |
| 8   | 0.00            | 0.00            | 0.00            | 0.00            |
| 9   | 0.00            | 0.00            | 0.07            | 0.11            |

The Value of the Cold, Collapsed Gas Fraction, \( f_{\text{cc}} \), at the End of the Simulations
direct impact on the dynamical evolution of the gas. Indeed, in the examples studied in this paper, our choice of rate coefficients has a great influence on the conclusion that we draw regarding whether or not the gas can cool and collapse within the lifetime of the protogalaxy (which will be of the order of a Hubble time). The presence of an ultraviolet background only serves to exacerbate this problem, as it increases the sensitivity of the predicted H$_2$ abundance to our choice of chemical rate coefficients.

Which of the two uncertainties we have discussed here is more important depends on the situation one is interested in. As mutual neutralization is important only in regions with a high fractional ionization, the uncertainty in the mutual neutralization rate coefficient will have a large impact on the outcome in such regions but will be of little to no importance in mostly neutral gas. Associative detachment, on the other hand, is important whenever one is interested in the gas-phase formation of H$_2$, which obviously encompasses a much wider range of situations, but the uncertainty in the associative detachment rate coefficient becomes important only when some other process, such as mutual neutralization or photodetachment, occurs fast enough to contend with it for the bulk of the H$_2$ ions.

Ultimately, the only way to remove these uncertainties will be to obtain more accurate rate coefficients for the relevant associative detachment and mutual neutralization processes. Most important for this will be new laboratory measurements for each reaction at cosmologically relevant collision energies supplemented by further theoretical calculations.
Stone, J. M., & Norman, M. L. 1992, ApJS, 80, 753
Susa, H., Uehara, H., Nishi, R., & Yamada, M. 1998, Prog. Theor. Phys., 100, 63
Szucs, S., Karemera, M., Terao, M., & Brouillard, F. 1984, J. Phys. B, 17, 1613
Tegmark, M., Silk, J., Rees, M., Blanchard, A., Abel, T., & Palla, F. 1997, ApJ, 474, 1
Tegmark, M., et al. 2004, Phys. Rev. D, 69, 103501
Thacker, R. J., Scannapieco, E., & Davis, M. 2002, ApJ, 581, 836
Wishart, A. W. 1979, MNRAS, 187, 59p
Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003, ApJ, 592, 645